

STIC-Biotech/ChemLib

164529

From: Whiteman, Brian
Sent: Wednesday, August 31, 2005 4:20 PM
To: STIC-Biotech/ChemLib
Subject: seq search

09/475,704 Barnett et al. 12/30/99

SEQ ID NO; 3 and 4
Please search protein and NA databases
1) the issued and published US application databases

Thank you,

Brian Whiteman
Remsen, 2D14
mail box 2C18
Patent Examiner - Art Unit 1635
United States Patent and Trademark Office
(571) 272-0764

8/31/05

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: 9/2/05
Date Completed: 9/2/05
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search reverse to
NA# 2+ AA# _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: DSF
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 20:59:32 ; Search time 261,349 Seconds
(without alignments)
9259.839 Million cell updates/sec

Title: US-09-475-704A-3
Perfect score: 1479
Sequence: 1 atgggagcgccgcgcagcat.....acgacccctgagccagtaa 1479

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCtUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1197.8	81.0	1515	4	US-09-475-515-4 Sequence 4, Appli
2	1197.8	81.0	4472	4	US-09-475-515-75 Sequence 75, Appli
3	1197.8	81.0	4608	4	US-09-475-515-76 Sequence 76, Appli
4	1197.8	81.0	4689	4	US-09-475-515-74 Sequence 74, Appli
5	1197.8	81.0	4766	4	US-09-475-515-73 Sequence 73, Appli
6	1195.8	80.9	2031	4	US-09-475-515-7 Sequence 7, Appli
7	1114.6	75.4	1853	4	US-09-475-515-5 Sequence 5, Appli
8	1114.6	75.4	1865	4	US-09-475-515-78 Sequence 78, Appli
9	1114.6	75.4	1865	4	US-09-475-515-79 Sequence 79, Appli
10	1114.6	75.4	4319	4	US-09-475-515-6 Sequence 6, Appli
11	1025.2	69.3	1268	4	US-09-475-515-9 Sequence 9, Appli
12	976.6	66.0	1521	4	US-09-952-060-27 Sequence 27, Appli
13	976.6	66.0	1532	4	US-09-974-702-1 Sequence 1, Appli
14	976.6	66.0	1532	4	US-09-818-443-1 Sequence 1, Appli
15	976.6	66.0	37474	4	US-09-952-060-25 Sequence 25, Appli
16	976.6	65.8	4053	4	US-09-952-060-34 Sequence 34, Appli
17	949.4	64.2	4307	4	US-09-936-572-2 Sequence 2, Appli
18	947.6	64.1	4307	4	US-09-552-950-2 Sequence 2, Appli
19	939.8	63.5	4327	4	US-09-936-572-14 Sequence 14, Appli
20	939.8	63.5	4353	4	US-09-936-572-13 Sequence 13, Appli
21	939.8	63.5	4642	4	US-09-936-572-12 Sequence 12, Appli
22	939.8	63.5	9772	4	US-09-552-950-5 Sequence 5, Appli
23	933	63.1	1482	4	US-09-818-443-4 Sequence 4, Appli
24	930	62.9	1479	4	US-09-952-060-32 Sequence 32, Appli
25	744.6	50.3	8366	4	US-09-872-733A-6 Sequence 6, Appli
26	738.2	49.9	1476	4	US-09-184-418C-100 Sequence 100, App
27	738.2	49.9	8959	4	US-09-184-418C-11 Sequence 11, Appli

28	733.4	49.6	4338	4	US-09-872-733A-1	Sequence 1, Appli
29	731.4	49.5	9913	4	US-09-827-688-11	Sequence 11, Appli
30	729.8	49.3	1485	4	US-09-184-418C-73	Sequence 73, Appli
31	729.8	49.3	9010	4	US-09-184-418C-8	Sequence 8, Appli
32	728.4	49.2	12523	4	US-09-991-258-1	Sequence 1, Appli
33	726	49.1	1476	4	US-09-991-258-4	Sequence 4, Appli
34	710.8	48.1	1486	4	US-09-184-418C-38	Sequence 38, Appli
35	710.8	48.1	8992	4	US-09-184-418C-4	Sequence 4, Appli
36	701.2	47.4	1496	4	US-09-184-418C-82	Sequence 82, Appli
37	701.2	47.4	8972	4	US-09-184-418C-9	Sequence 9, Appli
38	659.8	44.6	1503	4	US-09-184-418C-29	Sequence 29, Appli
39	659.8	44.6	8953	4	US-09-184-418C-3	Sequence 3, Appli
40	658.8	44.5	1479	4	US-09-184-418C-12	Sequence 12, Appli
41	658.8	44.5	8968	4	US-09-184-418C-1	Sequence 1, Appli
42	656.6	44.4	4307	4	US-09-184-418C-48	Sequence 48, Appli
43	656.6	44.4	8966	4	US-09-184-418C-5	Sequence 5, Appli
44	652.6	44.1	1494	4	US-09-184-418C-47	Sequence 47, Appli
45	650.2	44.0	4307	4	US-09-552-950-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-475-515-4
Sequence 4, Application US/09475515A
Patent No. 6602705
GENERAL INFORMATION:
APPLICANT: BARNETT, Susan
APPLICANT: ZUR MEGEDE, Jan
APPLICANT: SRIVASTAVA, Indresh
APPLICANT: LIAN, Ying
APPLICANT: HARTOG, Karin
APPLICANT: LIU, Hong
APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
APPLICANT: WALKER, Christopher
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
FILE REFERENCE: 1621.002
CURRENT APPLICATION NUMBER: US/09/475, 515A
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 90
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 1515
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-475-515-4

Query Match 81.0%; Score 1197.8; DB 4; Length 1515;
Best Local Similarity 89.3%; Pred. No. 9.3e-177;
Matches 1347; Conservative 0; Mismatches 132; Indels 30; Gaps 4;

QY	1	ATGGGGCGCCCGCGCAGCATCTGCGCGCGGCAAGCTGAGCGCTGGAGCGCATCCGC	60
DB	7	ATGGGGCGCCCGCGCAGCATCTGCGCGCGGCGAGCTGAGCAAGTGGAGAGATCCGC	66
QY	61	CTGGCGCCCGCGCGCAAGAGTGTACATGATGAAGCACCTGGTGGGCGCGCGAG	120
DB	67	CTGGCGCCCGCGCGCAAGAGTGTACATGATGAAGCACCTGGTGGGCGCGCGAG	126
QY	121	CTGGAGAGTTGCGCCCTGAACCCCGCGCTGCTGAGAGCAGCGAGGCTGCAAGCATC	180
DB	127	CTGGAGCGCTTGCCTGTGAACCCCGCGCTGCTGAGAGCAGCGAGGCTGCGCCAGATC	186
QY	181	ATCCGCAAGTGTGACCCCGCGCTGACAGCCGCGAGGAGAGCTGAAGAGCTGTCAAC	240
DB	187	CTGGCCAGCTGTGACCCCGCGCTGACAGCCGCGAGGAGAGCTGTGCGAGCTGTACAC	246

```
QY 241 ACCGTGGCCACCCTGTACTGCTGCAGAGAGATCGAGGTCCGCAACCAAGAGGCC 300
Db 247 ACCGTGGCCACCCTGTACTGCTGCAGAGAGATCGAGGTCCGCAACCAAGAGGCC 306
QY 301 CTGACAGAGATCGAGAGAGAGAGAGAGAGATCGAGAGATCGAGAGAGGCC 360
Db 307 CTGACAGAGATCGAGAGAGAGAGAGAGAGATCGAGAGAGAGAGAGAGGCC 366
QY 361 GCCG-----ACAAGGCAAGGTGAGCCAGAACTACCCCATCTGTGAGAACTG 408
Db 367 GCCGCGCGCAACCGGCAACAGCAGCCAGGTGAGCCAGAACTACCCCATCTGTGAGAACTG 426
QY 409 CAGGGCCAGATGTGTGACACAGGCCATGAGCCCGCAACCTGAAACGCTGGGTGAGGTG 468
Db 427 CAGGGCCAGATGTGTGACACAGGCCATGAGCCCGCAACCTGAAACGCTGGGTGAGGTG 486
QY 469 ATCGAGAGAGAGGCGCTTCAGCCCCCGAGGTGATCCCATGTTCAACCGCTGAGCGAGGC 528
Db 487 GTGAGAGAGAGAGGCGCTTCAGCCCCCGAGGTGATCCCATGTTCAACCGCTGAGCGAGGC 546
QY 529 GCCACCCCCCAGAGACCTTGAACACAGATGTTGAACAACCGTGGGGCGGCCACAGCGCCCATG 588
Db 547 GCCACCCCCCAGAGACCTTGAACACAGATGTTGAACAACCGTGGGGCGGCCACAGCGCCCATG 606
QY 589 CAGATGCTGAAGAGACACCATCAACGAGAGAGGCGCCGAGTGGGACCGCGGTGCAACCCCGTG 648
Db 607 CAGATGCTGAAGAGAGACCATCAACGAGAGAGGCGCCGAGTGGGACCGCGGTGCAACCCCGTG 666
QY 649 CAGCGCGCGCCCATCGCGCGCGCGCGCAAGTGGCGGAGCGCGCGAGACATCGCGCGG 708
Db 667 CAGCGCGCGCCCATCGCGCGCGCGCGCAAGTGGCGGAGCGCGCGAGACATCGCGCGG 726
QY 709 ACCACCAAGACCCCTGCAAGAGAGAGATGCTGATGACCAAGAACCCCGCATCCCGCTG 768
Db 727 ACCACCAAGACCCCTGCAAGAGAGAGATGCTGATGACCAAGAACCCCGCATCCCGCTG 786
QY 769 GGGGACATCTACAAGCGGTGATCATCTGGGCTGAAACAAGATCGTGGGATGTACAGC 828
Db 787 GGGGACATCTACAAGCGGTGATCATCTGGGCTGAAACAAGATCGTGGGATGTACAGC 846
QY 829 CCGGTGAGCATCTTGAATCAAGCAGAGGCCCAAGAGGCCCTTCCGCGACTACGTGGAC 888
Db 847 CCGGTGAGCATCTTGAATCAAGCAGAGGCCCAAGAGGCCCTTCCGCGACTACGTGGAC 906
QY 889 CGCTTCTTCAAGACCTTGGCGCGCGCGCAAGAGAGACCAAGAGGTGAAGAACTGATGAC 948
Db 907 CGCTTCTTCAAGACCTTGGCGCGCGCGCTGAGAGAGCAAGAGGTGAAGAACTGATGAC 966
QY 949 GACACCTTCTGTGTGAGAAACGCCAACCCCGACTGCAAGAACCATCTGCGCGCTCTCGGC 1008
Db 967 GACACCTTCTGTGTGAGAAACGCCAACCCCGACTGCAAGAACCATCTGAGAGGTCTCGGC 1026
QY 1009 CCGCGCGCGCGAGCTTGAAGAGAGATGATGACCGCTGCGAGGGCGTGGCGCGCGCCAGCGCAC 1068
Db 1027 CCGCGCGCGCGAGCTTGAAGAGAGATGATGACCGCTGCGAGGGCGTGGCGCGCGCCAGCGCAC 1086
QY 1069 AAGGCGCGCGTGTGGCGCGAGGCGGATGAGCCAGGCCAACCA-----CGAGCGTATGATG 1122
Db 1087 AAGGCGCGCGTGTGGCGCGAGGCGGATGAGCCAGGTGACGAACCCGCGAGCCATCATGATG 1146
QY 1123 CAGAAGAGCAACTTCAAGGGGCCCCCGCGCGCATGTCAGTGTCTTCAACTGCGGCAAGAG 1182
Db 1147 CAGCGCGGCAACTTCCGCAACAGCGGAAGACGTCAAGTGTCTTCAACTGCGGCAAGAG 1206
QY 1183 GGGCAGATCGCGCGCAATGCGCGCGCGCGCGCAAGAGGGCTGTGGAAGTGCAGGCAAG 1242
Db 1207 GGGCAGATCGCGCGCAATGCGCGCGCGCGCGCGCAAGAGGGCTGTGCGCGCGC 1266
QY 1243 GAGGGCCACAGATGAAGAGACTGCACGAGCGCCAGGCCAATTCTTGGGCAAGATCTGG 1302
Db 1267 GAGGGCCACAGATGAAGAGACTGCACGAGCGCCAGGCCAATTCTTGGGCAAGATCTGG 1326
QY 1303 CCGAGCCACAAGGCGCGCGCGCGCAACTTCTGTGAGAGCGCGCGCGAGCCACCGCGCGC 1362
```

```
Db 1327 CCGAGTACAAGGGCGCGCGCGCAACTTCTGTGAGAGCGCGCGCGAGCCACCGCGCGC 1386
QY 1363 CCGCGCGAGAGCTTCCGCTT-----GAGAGACCACCCCGCGCGAGAGAGAG-- 1413
Db 1387 CCGAGGAGAGCTTCCGCTTGGCGAGAGAGAACACACCCCGAGAGAGAGAGGCC 1446
QY 1414 --AGCAAGACCGCGAGACCTGACCGAGCTGAAAGAGCTGTTCGGCAACGACCCCTG 1470
Db 1447 ATGACAAGAGAGCTGTACCCCTTGACAGAGCTGCGAGAGCTGTTCGGCAACGACCCAGC 1506
QY 1471 AGCCAGTAA 1479
Db 1507 AGCCAGTAA 1515

RESULT 2
US-09-475-515-75
; Sequence 75, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 75
; LENGTH: 4472
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; US-09-475-515-75

Query Match 81.0%; Score 1197.8; DB 4; Length 4472;
Best Local Similarity 89.3%; Pred. No. 9,3e-177;
Matches 1347; Conservative 0; Mismatches 132; Indels 30; Gaps 4;

QY 1 ATGGGCGCGCGCGCGAGCATCTGCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCGC 60
Db 2939 ATGGGCGCGCGCGCGAGCGTGTGAGCGCGCGAGCTGACAAGTGGAGAGATCCGC 2998
QY 61 CTGCGCGCGCGCGCGCAAGAGTGTACATGATGAAGCACTGTGTGTGGCCAGCGCGAG 120
Db 2999 CTGCGCGCGCGCGCGCAAGAGAGTACAAAGCTGAAGCACATCTGTGTGGCCAGCGCGAG 3058
QY 121 CTGAGAAAGTTCGCGCTGAACCCCGCGCTGTGAGAGACCAAGAGGGCTGCAAGCATC 180
Db 3059 CTGAGAGCGTTTCGCGCTGAACCCCGCGCTGTGAGAGACCAAGAGGGCTGCGCCAGATC 3118
QY 181 ATCCGCGAGCTGACCCCGCGCTGACAGACCGGCAAGAGAGCTGAAGAGCTGTCAAC 240
Db 3119 CTGGGCGAGCTGACCCAGCTGTGACAGACCGGCAAGAGAGCTGCGAGCTGTACAAC 3178
QY 241 ACCGTGGCAACCTGTACTGCTGACAGAGAGATCGAGGTCCGCAACCAAGAGAGGCC 300
Db 3179 ACCGTGGCAACCTGTACTGCTGACAGAGAGATCGAGGTCCGCAACCAAGAGAGAGGCC 3238
QY 301 CTGACAAGATGAGAGAGAGAGAGAGAGATCGAGAGATCGAGAGAGATCGAGAGAGGCC 360
Db 3239 CTGAGAAGATGAGAGAGAGAGAGAGAGATCGAGAGAGAGAGAGAGAGAGAGAGAGGCC 3298
```



```
QY 361 GCCG-----ACAAAGGCAAGGTGAGCCAGAACTAACCCTGCTGCAGAACTG 408
Db 3299 GCCGCCGCGCACCGGCAACAGCAGCCAGCTGAGCCAGAACTAACCCTGCTGCAGAACTG 3358
QY 409 CAGGGCCAGATGTGTGCAACAGGCCATGAGCCCCCGACCTTGAACGCGCTGGTGAAGTG 468
Db 3359 CAGGGCCAGATGTGTGCAACAGGCCATGAGCCCCCGACCTTGAACGCGCTGGTGAAGTG 3418
QY 469 ATCGAGAGAAAGGCTTTCAGCCCCGAGGTGATCCCAATGTTCAACCGCCTGAGCGAGGCG 528
Db 3419 GTGAGAGAGAAAGGCTTTCAGCCCCGAGGTGATCCCAATGTTCAACCGCCTGAGCGAGGCG 3478
QY 529 GCCACCCCCCAGGACCTGAAACAGATGTTGAACACCGTGGGCGGCAACGAGCCGCCATG 588
Db 3479 GCCACCCCCCAGGACCTGAAACAGATGTTGAACACCGTGGGCGGCAACGAGCCGCCATG 3538
QY 589 CAGATGCTGAAGAGACACCATCAACGAGAGGCGCGGAGTGGGACCGCGTGCAACCCCGTG 648
Db 3539 CAGATGCTGAAGAGAGACCATCAACGAGAGGCGCGGAGTGGGACCGCGTGCAACCCCGTG 3598
QY 649 CAGCGCGGCCCCCATCGCCCCCGGCGCAAGTGGCGGAGCCCGGCAACGACATCGCGCGC 708
Db 3599 CAGCGCGGCCCCCATCGCCCCCGGCGCAAGTGGCGGAGCCCGGCAACGACATCGCGCGC 3658
QY 709 ACCACCAACACCCCTGCAAGAGAGAGATCGCTGATGACCAAGCAACCCCGCATCCCGCTG 768
Db 3659 ACCACCAACACCCCTGCAAGAGAGAGATCGCTGATGACCAAGCAACCCCGCATCCCGCTG 3718
QY 769 GGCGCATCTACAAGCGGTGATCATCTGGGCTGAAACAAGATCGTGGGATGTACAGC 828
Db 3719 GGCGCATCTACAAGCGGTGATCATCTGGGCTGAAACAAGATCGTGGGATGTACAGC 3778
QY 829 CCGGTGAGCATCTCTGATCATCAAGCAGAGGCCCCCAAGAGCCCTTCCCGCACTACGTGAC 888
Db 3779 CCGGTGAGCATCTCTGATCATCAAGCAGAGGCCCCCAAGAGCCCTTCCCGCACTACGTGAC 3838
QY 889 CGCTTCTTCAAGACCTCTGCGCGCGCGGAGAGAGACCAACGAGGTGAAGAACTGATGACC 948
Db 3839 CGCTTCTTCAAGACCTCTGCGCGCGCGTGAAGAGAGACCAACGAGGTGAAGAACTGATGACC 3898
QY 949 GACACCTCTGCTGTGTCAGAAACGCCAACCCCGACTGCAAGAACCATCTGCGCGCTCTGCGC 1008
Db 3899 GACACCTCTGCTGTGTCAGAAACGCCAACCCCGACTGCAAGAACCATCTGCGCGCTCTGCGC 3958
QY 1009 CCGCGCGCGCAGCCTGAGAGAGATGATGACCGCTGCGAGGGCGTGGCGGCCCCCAGCAC 1068
Db 3959 CCGCGCGCGCAGCCTGAGAGAGATGATGACCGCTGCGAGGGCGTGGCGGCCCCCAGCAC 4018
QY 1069 AAGGCCGCGCTGTGCGCGGAGGCGATGAGCCAGGCCAAACA-----CGAGCGTATGATG 1122
Db 4019 AAGGCCGCGCTGTGCGCGGAGGCGATGAGCCAGGTGACGAACCCCGGAGCATCATGATG 4078
QY 1123 CAGAAAGACAACCTTCAAGGGCCCCCGCGCATCTGTCAAGTGTCTTCACTGCGGCAAGAG 1182
Db 4079 CAGCGCGGCAACTTTCGCAACAGCGGAAGACCGTCAAGTGTCTTCACTGCGGCAAGAG 4138
QY 1183 GGCCACATGCGCGCGCAACTGCGCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGCAGGCAAG 1242
Db 4139 GGCCACATGCGCGCGCAACTGCGCGCGCCCCCGCAAGAGGGCTGCTGCGCGCTGCGCGC 4198
QY 1243 GAGGGCCACCAAGATGAAGAACTGACCGAGCGCCAGGCCAACTTCTGCGGCAAGATCTGG 1302
Db 4199 GAGGGCCACCAAGATGAAGAACTGACCGAGCGCCAGGCCAACTTCTGCGGCAAGATCTGG 4258
QY 1303 CCGAGCCACAAGGCGCGCCCCCGCAACTTCTGCAAGGCGCGCCCCCGAGCCACCGCGCCC 1362
Db 4259 CCGAGCTACAAGGCGCGCCCCCGCAACTTCTGCAAGGCGCGCCCCCGAGCCACCGCGCCC 4318
QY 1363 CCGCGGAGAGCTTCCGCTTC-----GAGGAGACCAACCCCGGCGCAAGCAGAGAG-- 1413
Db 4319 CCGGAGAGAGCTTCCGCTTCGCGCGAGGAGAAAGACCAACCCCGAGCCGAAGCAGAGGCC 4378
```

```
QY 1414 ---AGCAAGAGCCGCGAGACCCCTGACCAAGCCTGTGTTGGCAACGACCCCTG 1470
Db 4379 ATCGACAAGAGAGCTGTACCCCTGACCAAGCCTGCGCAAGCCTGTTGGCAACGACCCAGC 4438
QY 1471 AGCCAGTAA 1479
Db 4439 AGCCAGTAA 4447

RESULT 3
US-09-475-515-76
; Sequence 76, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 76
; LENGTH: 4608
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; US-09-475-515-76

Query Match      81.0%; Score 1197.8; DB 4; Length 4608;
Best Local Similarity 89.3%; Pred. No. 9.3e-177;
Matches 1347; Conservative 0; Mismatches 132; Indels 30; Gaps 4;

QY 1 ATGGGCGCGCGCGCAGCATCTGCGCGCGGCGCAAGCTGAGCGCTGGAGCGCATCCGC 60
Db 3075 ATGGGCGCGCGCGCAGCATCTGCGCGCGGCGCAAGCTGAGCGCGCTGGAGCGCATCCGC 3134
QY 61 CTGCGCGCGCGCGCAAGAGTCTACATGATGAAGACCTGTGTGGGCCAGCGCGAG 120
Db 3135 CTGCGCGCGCGCGCAAGAGTCTACATGATGAAGACACATGTGTGGGCCAGCGCGAG 3194
QY 121 CTGGAAGATTGCGCCCTGAACCCCGGCTGTGAGAGACCAAGCGGCTGCAACGATC 180
Db 3195 CTGGAAGCTTGCCTGTGAACCCCGGCTGTGAGAGACCAAGCGGCTGCGCGCATC 3254
QY 181 ATCCGCGAGCTGCAACCCGCTGCAAGACCGGCAAGAGAGAGAGCTGAAGAGCTGTTCAAC 240
Db 3255 CTGGGCGAGCTGCAACCCGCTGCAAGACCGGCAAGAGAGAGAGCTGCGAGCTGTACAAC 3314
QY 241 ACCGTGGCAACCTGTACTGCGTGCACGAGAAAGATGAGGTCCGCGCAACCAAGAGGGCC 300
Db 3315 ACCGTGGCAACCTGTACTGCGTGCACGAGCGCATGACGTCAAGAGACCAAGAGAGGGCC 3374
QY 301 CTGGAACAAGATGAGAGAGAGCAACAAGTGCACAGAGAAAGATCCAGAGCGGAGGCC 360
Db 3375 CTGGAAGAAGATGAGAGAGAGCAACAAGTGCACAGAGAAAGAGCCAGAGCGCGCGCC 3434
QY 361 GCCG-----ACAAAGGCAAGGTGAGCCAGAACTAACCCTGCTGCAGAACTG 408
Db 3435 GCCGCCGCGCACCGGCAACAGCAGCCAGCTGAGCCAGAACTAACCCTGCTGCAGAACTG 3494
QY 409 CAGGGCCAGATGTGTGCAACAGGCCATGAGCCCCCGACCTTGAACGCGCTGGTGAAGTG 468
```

```
Db 3495 CAGGGCCAGATGTGCAACAGGCCATCAGCCCCCGCACCTGAACGCTGGGTGAAGTG 3554
QY 469 ATGAGAGAGAGGCTTCAGCCCCCGAGGTGATCCCATGTTCAACCGCCTGAGCGAGGC 528
Db 3555 GTGAGAGAGAGGCTTCAGCCCCCGAGGTGATCCCATGTTCAACCGCCTGAGCGAGGC 3614
QY 529 GCCACCCCCCAGACTGAACACGATGTTGAACACCGTGGGCGGCCACAGGCCGCATG 588
Db 3615 GCCACCCCCCAGACTGAACACGATGTTGAACACCGTGGGCGGCCACAGGCCGCATG 3674
QY 589 CAGATGCTGAAGGACACCATCAACGAGAGAGGCCCGAGTGGGACCCCGTGCACCCCGTG 648
Db 3675 CAGATGCTGAAGGACCATCAACGAGAGAGGCCCGAGTGGGACCCCGTGCACCCCGTG 3734
QY 649 CACGCGGGCCCCCATCGCCCCCGCGCATGCGCGAGCCCCCGCGCAGACATCGCGGC 708
Db 3735 CACGCGGGCCCCCATCGCCCCCGCGCATGCGCGAGCCCCCGCGCAGACATCGCGGC 3794
QY 709 ACCACCAAGCACCCTGAGGAGGAGATGCTGATGACCAAGCAACCCGCCATCCCGGTG 768
Db 3795 ACCACCAAGCACCCTGAGGAGGAGATGCTGATGACCAACCAACCCGCCATCCCGGTG 3854
QY 769 GGGCACAATCTACAAGCGGTGATCATCTGGGCTTGAACAAGATCGTGGGATGACAGC 828
Db 3855 GGGCAGATCTACAAGCGGTGATCATCTGGGCTTGAACAAGATCGTGGGATGACAGC 3914
QY 829 CCGGTGAGCATCCTGAGCATCAAGCAGGGGCCCAAGAGGCCCTTCGCGACTACGTGGAC 888
Db 3915 CCCACCAAGCATCCTGAGCATCGCCAGGGGCCCAAGAGGCCCTTCGCGACTACGTGGAC 3974
QY 889 CGCTTCTTCAAGACCCCTGCGCGCGCGCAGAGAGCAACCAAGAGGTGAAGAACTGATGACC 948
Db 3975 CGCTTCTTCAAGACCCCTGCGCGCGCTAGCAGGCCAGCCAGACGTGAAGAACTGATGACC 4034
QY 949 GACACCTCTGCTGTGTCAGAAACGCCAACCCCGACTGCAAGACCATCTGCGGCTCTGCGC 1008
Db 4035 GAGACCTCTGCTGTGTCAGAAACGCCAACCCCGACTGCAAGACCATCTGAAAGGCTCTGCGC 4094
QY 1009 CCGCGGCGCCAGCCTGAGGAGATGATGACCGCCTGCCAGGGCGGTGGCGGCCAGCCAC 1068
Db 4095 CCGCGGCGCCACCCTGAGGAGATGATGACCGCCTGCCAGGGCGGTGGCGGCCAGCCAC 4154
QY 1069 AAGGCGCGCGTGTGGCGCGAGCGATGAGCCAGGCCACA-----CGAGCGTATGATG 1122
Db 4155 AAGGCGCGCGTGTGGCGCGAGCGATGAGCCAGGTGACGAACCCGCGGACCATCATGATG 4214
QY 1123 CAGAAAGACCACTTCAAGGGCCCCCGCGCATCTGTCAAGTGTCTTCAACTGCGGCAAGAG 1182
Db 4215 CAGCGCGGCAACTTCCGCAACCAAGCGGAAGACCGTCAAGTGTCTTCAACTGCGGCAAGAG 4274
QY 1183 GGCCACATGCGCCCGCAACTGCGCGCGCCCCCGCAAGAGGGCTGTGAAAGTGGCGCAAG 1242
Db 4275 GGCCACACCGCCAGGAATGCGCGCGCCCCCGCAAGAGGGCTGTGCGCTGCGCGCGC 4334
QY 1243 GAGGGCCACAGATGAAGGACTGACCGAGCGCCAGGCCAATTCTCTGGGCAAGATCTGG 1302
Db 4335 GAGGGCCACAGATGAAGGACTGACCGAGCGCCAGGCCAATTCTCTGGGCAAGATCTGG 4394
QY 1303 CCCAGCCACAAGGGCGCGCCCCCGCAACTTCTGCAAGCGCGCCGAGCCACCGCGCCC 1362
Db 4395 CCCAGCTTACAAGGGCGCGCCCCCGCAACTTCTGCAAGCGCGCCGAGCCACCGCGCCC 4454
QY 1363 CCGCGGAGAGCTTCCGCTTC-----GAGGAGACCAACCCCGCGCAGAAAGCAGAG--- 1413
Db 4455 CCGGAGAGAGCTTCCGCTTCGCGAGGAGAAAGACCAACCCAGCCAGAAAGCAGAGCCC 4514
QY 1414 ---AGCAAGGAGCCGCGAGACCCTGACAGCCTGAAGAGCCTGTTGGCAAGAGACCCCTG 1470
Db 4515 ATCGACAAGAGAGCTGTATCCCTGACAGCCTGCGCAGCTGTTCGCAACGACCCCAAGC 4574
QY 1471 AGCCAGTAA 1479
Db 4575 AGCCAGTAA 4583
```

```
RESULT 4
US-09-475-515-74
; Sequence 74, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475, 515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 74
; LENGTH: 4689
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; US-09-475-515-74

Query Match      81.0%; Score 1197.8; DB 4; Length 4689;
Best Local Similarity 89.3%; Pred. No. 9.3e-177;
Matches 1347; Conservative 0; Mismatches 132; Indels 30; Gaps 4;

QY 1 ATGGGCGCGCGCGCCAGCATCTGCGGCGGCGCAAGCTGACGCTGGGAGCGCATCCGC 60
Db 3156 ATGGGCGCGCGCGCCAGCGTGTGAGCGGCGCGGAGCTGACAACTGGGAGATCCGC 3215
QY 61 CTGCGCGCGCGCGCGCAAGAGTGTACATATGAAGCACTTGGTGTGGCCAGCGCGAG 120
Db 3216 CTGCGCGCGCGCGCGCAAGAGTGAAGCTGAAGCACTGTGTGGCCAGCGCGAG 3275
QY 121 CTGGAAGATGTGCGCTGAAACCCCGCGCTGTGAGACCAAGGAGGCTGCAAGATC 180
Db 3276 CTGGAAGCTTGTGCGGTGAACCCCGCGCTGTGAGACCAAGGAGGCTGCGCCAGATC 3335
QY 181 ATCCGCAGCTGCACCCCGCGCTGCAGACCGCGCAGGAGCTGAAGAGCTGTTCAC 240
Db 3336 CTGGGCGAGCTGCAGCCAGCTGCAGACCGCGCAGGAGCTGCGCAGCTGTACAC 3395
QY 241 ACCGTGGCCACCTGTACTGCGTGCACGAGAAATCGAGTCCGCGACCAAGAGAGCC 300
Db 3396 ACCGTGGCCACCTGTACTGCGTGCACCAAGCATGACGTCAAGACCAAGAGAGCC 3455
QY 301 CTGGAACAAGATGAGAGGAGCAACAAGTGCAGACAGATCCAGAGCCGAGGCC 360
Db 3456 CTGGAAGAAGATGAGAGGAGCAACAAGTCCAAAGAAAGGCCAGCAGCGCGCGCC 3515
QY 361 GCGC-----ACAAAGGCAAGGTGAGCCAGAACTTACCCCATGCTGAGAACTG 408
Db 3516 GCGCGCGGCAACGCAACAGCAGCCAGGTGAGCCAGAACTTACCCCATGCTGAGAACTG 3575
QY 409 CAGGGCCAGATGTGCAACGAGCCATCAGCCCGCGCAGCTGAAAGCCTGGTGAAGTG 468
Db 3576 CAGGGCCAGATGTGCAACGAGCCATCAGCCCGCGCAGCTGAAAGCCTGGTGAAGTG 3635
QY 469 ATCGAGAGAGAGCCTTCAGCCCGAGGTGATCCCATGTTCAACGCGCTGAGCAGAGGC 528
Db 3636 GTGGAAGAGAGAGCCTTCAGCCCGAGGTGATCCCATGTTCAACGCGCTGAGCAGAGGC 3695
QY 529 GCCACCCCCCAGACTGAACACGATGTTGAACACCGTGGGCGGCCACAGGCCGCATG 588
```

```
Db      3696  GCGACCCCGGAGGAGCTGAGACGATGTTGAACACCGTGGCGGCGCCACGAGCGCGCATG 3755
Qy      589   CAGATGCTGAAGAGACACCATCAACGAGAGCGCGCGGAGTGGGACCGCTGCACCCCGTG 648
Db      3756  CAGATGCTGAAGAGAGACCATCAACGAGAGCGCGCGGAGTGGGACCGCTGCACCCCGTG 3815
Qy      649   CAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 708
Db      3816  CAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3875
Qy      709   ACCACGAGACCCCTGCAAGAGCAGATCGCTGGATGACCAACCCCGCGCATCCCGCTG 768
Db      3876  ACCACGAGACCCCTGCAAGAGCAGATCGCTGGATGACCAACCCCGCGCATCCCGCTG 3935
Qy      769   GCGGACATCTACAAGCGGTGGATCATCTGGGCGCTGAACAAGATCGTGGCGATGTACAG 828
Db      3936  GCGGAGATCTACAAGCGGTGGATCATCTGGGCGCTGAACAAGATCGTGGCGATGTACAG 3995
Qy      829   CCGGTGAGCATCTGAGCATCAAGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 888
Db      3996  CCGACGAGCATCTGAGCATCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4055
Qy      889   CGCTTCTTCAAGACCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 948
Db      4056  CGCTTCTTCAAGACCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4115
Qy      949   GACACCTGCTGTGTGCAAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTGCG 1008
Db      4116  GAGACCTGCTGTGTGCAAGAACGCCAACCCCGACTGCAAGACCATCTGAGAGCTTGG 4175
Qy      1009  CCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1068
Db      4176  CCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4235
Qy      1069  AAGCGCGCGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1122
Db      4236  AAGCGCGCGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4295
Qy      1123  CAGAAGAGCAACTTCAAGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1182
Db      4296  CAGCGCGCGCAACTTCCGCAACAGCGGAGAACCGTCAAGTGTCTTCAACTGCGCG 4355
Qy      1183  GCGCACATGCGCGCGCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1242
Db      4356  GCGCACACCGCGCGAGAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4415
Qy      1243  GAGGCGCACAGATGAAGACTGCAACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1302
Db      4416  GAGGCGCACAGATGAAGACTGCAACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4475
Qy      1303  CCGAGCGCAAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1362
Db      4476  CCGAGCTTCAAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4535
Qy      1363  CCGCGGAGAGCTTCCGCTTC-----GAGGAGACCAACCGCGCGCGCGCGCGCGCG 1413
Db      4536  CCGGAGAGAGCTTCCGCTTCGCGCGAGAGAGAACCAACCGCGCGCGCGCGCGCGCG 4595
Qy      1414  ---AGCAAGAGACCGCGAGACCTGACCAAGAGCTGTGGCAAGAGAGAGAG--- 1470
Db      4596  ATGCAAGAGAGCTGTATCCCTGACCAAGCTGCGGAGCTGTGGCAAGAGAGAGAG 4655
Qy      1471  AGCCAGTAA 1479
Db      4656  AGCCAGTAA 4664
```

RESULT 5
US-09-475-515-73
; Sequence 73, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:

```
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 73
; LENGTH: 4766
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160.modUS4.gag.modSF2
US-09-475-515-73
```

```
Query Match      81.0%; Score 1197.8; DB 4; Length 4766;
Best Local Similarity 89.3%; Pred. No. 9.3e-177;
Matches 1347; Conservative 0; Mismatches 132; Indels 30; Gaps 4;

Qy      1  ATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
Db      3233  ATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3292
Qy      61  CTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db      3293  CTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3352
Qy      121  CTGGAAGATTGCGCTTGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Db      3353  CTGGAAGATTGCGCTTGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3412
Qy      181  ATCCGCGAGCTGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Db      3413  CTGGCGCAGCTGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3472
Qy      241  ACCGTGGCCACCTGTACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Db      3473  ACCGTGGCCACCTGTACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3532
Qy      301  CTGGAAGATCGAGAGAGAGCAAGTGCAGCAGAGAGATCCAGAGAGAGAGAGAGAG 360
Db      3533  CTGGAAGATCGAGAGAGAGCAAGTGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAG 3592
Qy      361  GCCG-----ACAAGGCAAGTGAAGCGCAAGTACCCCATGCTGCAAGAACTG 408
Db      3593  GCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3652
Qy      409  CAGGCGCAGATGTGCAACGAGCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 468
Db      3653  CAGGCGCAGATGTGCAACGAGCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3712
Qy      469  ATCGAGAGAGAGGCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 528
Db      3713  GTGAGAGAGAGAGGCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3772
Qy      529  GCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 588
Db      3773  GCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3832
Qy      589  CAGATGCTGAAGAGACCATCAACGAGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCG 648
Db      3833  CAGATGCTGAAGAGACCATCAACGAGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCG 3892
```


QY	769	GGCGACATCTACAAGCGGTGATCATCTTGCGCTGAACAAGATCGTGGATGTACAGC	828
Db	787	GGCGAGATCTACAAGCGGTGATCATCTTGCGCTGAACAAGATCGTGGATGTACAGC	846
QY	829	CCCGTGAGCATCTTGACATCAAGAGGGCCCAAGAGCCCTTCGCGACTACGTGAC	888
Db	847	CCCAACAGCATCTTGACATCCGCGAGGGCCCAAGAGCCCTTCGCGACTACGTGAC	906
QY	889	CGCTTCTTCAAGACCCCTGCGCGCGAGCAGAGCAACCAGAGGTGAAGAACTGATGACC	948
Db	907	CGCTTCTTCAAGACCCCTGCGCGCGAGCAGAGCAACCAGAGGTGAAGAACTGATGACC	966
QY	949	GACACCCCTGCTGTGTGAGAACGCCAACCCCGACTGCAAGAACCTCTGCGCTCTGCGC	1008
Db	967	GAGACCCCTGCTGTGTGAGAACGCCAACCCCGACTGCAAGAACCTCTGCGCTCTGCGC	1026
QY	1009	CCCGGCGCCAGCCTTGGAGAGATGATGACCGCTGCGCAGGGCGTGGCGGCCAGCCAC	1068
Db	1027	CCCGGCGCCAGCCTTGGAGAGATGATGACCGCTGCGCAGGGCGTGGCGGCCAGCCAC	1086
QY	1069	AAGGCGCGCGTGTGTGCGCGAGGCGATGAGCCAGGCCAACACACACACACACACACAC	1122
Db	1087	AAGGCGCGCGTGTGTGCGCGAGGCGATGAGCCAGGCCAACACACACACACACACACAC	1146
QY	1123	CAGAGAGCAACTTCAAGGGCCCCCGCGCATCTGCAAGTGTCTCACTGCGCAAGAG	1182
Db	1147	CAGCGCGCAACTTCCGCAACCAAGCGAAGACCGTCAAGTGTCTCACTGCGCAAGAG	1206
QY	1183	GGCCACATCGCCCGCAACTGCGCGCGCCCCCGCAAGAGGGCTGTGMACTGCGCGAAG	1242
Db	1207	GGCCACATCGCCCGCAACTGCGCGCGCCCCCGCAAGAGGGCTGTGMACTGCGCGAAG	1266
QY	1243	GAGGGCCACAGATGAAGGACTGCAACCGAGCGCGCAACTCTCTGGGCAAGATCTGG	1302
Db	1267	GAGGGCCACAGATGAAGGACTGCAACCGAGCGCGCAACTCTCTGGGCAAGATCTGG	1326
QY	1303	CCCAAGCCACAAGGGCGCGCCCGGCAACTCTCTGAGAGCGCGCGAGCCCAAGCCCGCC	1362
Db	1327	CCCAAGCCACAAGGGCGCGCCCGGCAACTCTCTGAGAGCGCGCGAGCCCAAGCCCGCC	1386
QY	1363	CCCGCCGAGAGCTTCCGCTTC-----GAGGAGACCAACCCCGCGCGAGAGCAGAG--	1413
Db	1387	CCCGAGGAGAGCTTCCGCTTCGCGGAGAGGAGAACCAACCCCGAGCGAGAGAGAGCC	1446
QY	1414	---AGCAAGGAGCGCGAGACCTTGACCAAGCCTTGAAAGAGCCTGTGGCAACGACCCCTG	1470
Db	1447	ATCGACAAAGAGCTGTACCCCTTGACCAAGCCTTGCGAGCCTGTGGCAACGACCCCGAGC	1506
QY	1471	AGCCAGT 1477	
Db	1507	AGCCAGT 1513	
RESULT 7			
US-09-475-515-5			
; Sequence 5, Application US/09475515A			
; Patent No. 6602705			
; GENERAL INFORMATION:			
; APPLICANT: BARNETT, Susan			
; APPLICANT: ZUR MEGEDE, Jan			
; APPLICANT: SRIVASTAVA, Indresh			
; APPLICANT: LIAN, Ying			
; APPLICANT: HARTOG, Karin			
; APPLICANT: LIU, Hong			
; APPLICANT: GREER, Catherine			
; APPLICANT: SELBY, Mark			
; APPLICANT: WALKER, Christopher			
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION			
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES			
; FILE REFERENCE: 1621.002			
; CURRENT APPLICATION NUMBER: US/09/475, 515A			
; CURRENT FILING DATE: 1999-12-30			

```

; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1853
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: HIV-Gag-protease
US-09-475-515-5

Query Match              75.4%; Score 1114.6; DB 4; Length 1853;
Best Local Similarity    85.8%; Pred. No. 6.6e-164;
Matches 1295; Conservative 0; Mismatches 184; Indels 30; Gaps 4;

QY      1 ATGGGCGCCCGCGCCAGCATCCTGCGCGGCAAGCTGAAGCCTGGAGCGCATCCGC 60
      |||||
Db       7 ATGGGCGCCCGCGCCAGCGTGTCTGAGCGCGCGAGCTGACAAAGTGGAAGATCCGC 66
      |||||

QY      61 CTGGCGCCCGCGCGCAAGAGTCTATCATGATGAAGCACTGTGTGTGGGCCCGCGAG 120
      |||||
Db       67 CTGGCGCCCGCGCGCAAGAAAGTACAAAGCTGAAGCACTGTGTGGGCCCGCGAG 126
      |||||

QY      121 CTGGAAGAAGTTGCGCCCTGAACCCCGGCTGTCTGGAAGACCAGGAGGGCTGCAAGCATC 180
      |||||
Db       127 CTGGAAGCGCTTGCGCCGTGAACCCCGGCTGTCTGGAAGACCAGGAGGGCTGCGCAGATC 186
      |||||

QY      181 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGCGAGGAGAGCTGAAGAGCCTGTTCAAC 240
      |||||
Db       187 CTGGGCGCAGCTGCAGCCCAAGCCTGCAGACCGGCGAGGAGCTGCGCAGCCTGTACAAAC 246
      |||||

QY      241 ACCGTGGCCACCCCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACCAAGAGAGGCC 300
      |||||
Db       247 ACCGTGGCCACCCCTGTACTGCGTGCACCAAGCGCATCGACGTCAAGGACCAAGAGAGGCC 306
      |||||

QY      301 CTGGAACAAGATCGAGGAGGAGCAAAAGTCCAGCAAGAATCCAGCAGGCCGAGGCC 360
      |||||
Db       307 CTGGAAGAAGATCGAGGAGGAGCAAAAGTCCAAAGAAAGGCCCAAGCAGGCCCGCGCC 366
      |||||

QY      361 GCCG-----ACAAAGGCGAAGGTGAGCCAGAACTAACCCCATCGTGCAGAACTG 408
      |||||
Db       367 GCCGCGCGGCAACCGGCAACAGCAGCCAGGTGAGCCAGAACTAACCCCATCGTGCAGAACTG 426
      |||||

QY      409 CAGGGCCAGATGCTGCACCCAGGCCATCAGCCCCCGCACCCCTGAAACGCTGGGTGAAGTG 468
      |||||
Db       427 CAGGGCCAGATGCTGCACCCAGGCCATCAGCCCCCGCACCCCTGAAACGCTGGGTGAAGTG 486
      |||||

QY      469 ATCGAGGAGAAAGGCTTACAGCCCCGAGGTGATCCCATGTTCAACGCGCCTGAGCGAGGGC 528
      |||||
Db       487 GTGGAAGAGAAAGGCTTACAGCCCCGAGGTGATCCCATGTTCAACGCGCCTGAGCGAGGGC 546
      |||||

QY      529 GCCAACCCTCCAGGACCTGAACACGATGTTGAACACCGTGGGCGGCCAACAGGCCCGCCATG 588
      |||||
Db       547 GCCAACCCTCCAGGACCTGAACACGATGTTGAACACCGTGGGCGGCCAACAGGCCCGCCATG 606
      |||||

QY      589 CAGATGCTGAAGGAGCAACCATCAACGAGGAGGCGCGGAGTGGGACCGCGTGCACCCCGTG 648
      |||||
Db       607 CAGATGCTGAAGGAGCAACCATCAACGAGGAGGCGCGGAGTGGGACCGCGTGCACCCCGTG 666
      |||||

QY      649 CAGCGCGGCGCCATCGCCCCCGGCGCAGATGCGCGAGCCCGCGGAGCGACATCGCGCGGC 708
      |||||
Db       667 CAGCGCGGCGCCATCGCCCCCGGCGCAGATGCGCGAGCCCGCGGAGCGACATCGCGCGGC 726
      |||||

QY      709 ACCACCAAGCACCTCTGCAAGAGCAGATCGCCTGGATGACCAACAACCCCCCATCCCGGTG 768
      |||||
Db       727 ACCACCAAGCACCTCTGCAAGAGCAGATCGCCTGGATGACCAACAACCCCCCATCCCGGTG 786
      |||||

QY      769 GGCACACATCTAACAGCGGTGATCATCTGGGGCCTGAACAAGATCGTGGGATGTACAGC 828
      |||||
Db       787 GGCACACATCTAACAGCGGTGATCATCTGGGGCCTGAACAAGATCGTGGGATGTACAGC 846
      |||||

QY      829 CCCGTGAGCATCTTGACATCAAGCAGGGGCCCAAGGAGCCTTCCGCGCACTACGTGAC 888
      |||||

```

```
Db      847  CCACACGATCCTGGACATCCGCCAGGGCCCCAAGAGCCCTTCCGCGACTACGTGAC  906
QY      889  CGCTTCTTCAAGACCTGCGCGCCGAGACAGACCCAGAGGTGAAGACTGATGACC  948
Db      907  CGCTTCTACAAGACCTGCGCGCTGAGCAGGCCAGCCAGGACGTGAAGACTGATGACC  966
QY      949  GACACCTGCTGTGAGAACGCCAACCCCGACTGCAAGACCATCTGCGGCTCTGGC  1008
Db      967  GAGACCTGCTGTGAGAACGCCAACCCCGACTGCAAGACCATCTGAGGCTCTGGC  1026
QY      1009  CCGCGCGCCAGCTGAGAGATGATGACCGCTGCCAGGGCGTGGCGGCCCAAGCAC  1068
Db      1027  CCGCGCGCCAGCTGAGAGATGATGACCGCTGCCAGGGCGTGGCGGCCCAAGCAC  1086
QY      1069  AAGCGCGCGTGTGGCGGAGCGATGAGCCAGCCACA-----CCAGCGTATGATG  1122
Db      1087  AAGCGCGCGTGTGGCGGAGCGATGAGCCAGCGAACCCCGGAGACCATCATGATG  1146
QY      1123  CAGAGAGCAACTTCAAGGGCCCCCGCGCATCGTCAAGTGTCTCACTGGCGCAAGAG  1182
Db      1147  CAGCGCGGCAACTTCCGCAACCAAGCGAAGAACCGTCAAGTGTCTCACTGGCGCAAGAG  1206
QY      1183  GGCCACATCGCCCGCAACTGCGCGCGCCCCCGCAAGAGGGCTGTGGAAGTGGCGCAAG  1242
Db      1207  GGCCACACCGCCAGGAATGCGCGCGCCCCCGCAAGAGGGCTGTGCGCGCGCGC  1266
QY      1243  GAGGGCCACCATGTAAGGACTGCACCGAGCGCCAGCCCACTTCTGGGCAAGATCTGG  1302
Db      1267  GAAGGACACCAATGAAAGATTGCACTGAGAGACAGGCTAATTTTAAAGGAAGATCTGG  1326
QY      1303  CCCAGCCACAAGGGCGCGCCCGGCACTTCTGCAGAGCGCGCCGAGCCCAAGCCGCCCC  1362
Db      1327  CCTTCTACAAGGGAAGGCCAGGAAATTTTCTTCAAGAGCAGACCAAGCCCAAGCCCCA  1386
QY      1363  CCGCGCGAGAGCTTCCGCTTGAAGAG-----ACCACCCCCGGCGCAAGAGCAGAG---  1413
Db      1387  CCAGAAGAGAGCTTCAAGTTTGGGAGAGAGAAAACAATCCTCCTCAGAAAGCAGAGCCG  1446
QY      1414  ---AGCAAGGACCGCGAGACCTGACCAAGCCTGAAGAGCCTGTTCCGCAACGACCCCTG  1470
Db      1447  ATAGACAAAGAACTGTATCCTTAACTTCCCTCAGATCACTCTTTGGCAACGACCCCTCG  1506
QY      1471  AGCCAGTAA 1479
Db      1507  TCACAGTAA 1515
```

```
RESULT 8
US-09-475-515-78
; Sequence 78, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MESEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 78
; LENGTH: 1865
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
; OTHER INFORMATION: Description of Artificial Sequence: GPI
US-09-475-515-78
```

```
Query Match      75.4%; Score 1114.6; DB 4; Length 1865;
Best Local Similarity 85.8%; Pred. No. 6.6e-164;
Matches 1295; Conservative 0; Mismatches 184; Indels 30; Gaps 4;
```

```
QY      1  ATGGGCGCGCGCCAGCATCTGCGCGCGCAAGCTGAGCGCCTGGAGCGCATCCGC  60
Db      13  ATGGGCGCGCGCCAGCGTGTGAGCGCGCGCAGCTGGAACAGTGGAGAAATCCGC  72
QY      61  CTGCGCGCGCGCGCAAGAGTGTACATGATGAAGCACCTGTGTGGCGCAGCGGAG  120
Db      73  CTGCGCGCGCGCGCAAGAAAGTACAAAGCTGAAGCACATGTGTGGCGCAGCGGAG  132
QY      121  CTGGAAGATTGCGCCTGAACCCCGCCTGTGGAAGACCAAGAGGCTGCAAGCAGATC  180
Db      133  CTGGAAGCTTTCGCGGTGAACCCCGCCTGTGGAAGACCAAGAGGCTGCGCCAGATC  192
QY      181  ATCCGCGAGCTGACACCCCGCCTGCAGACCGGCAAGAGAGCTGAAGAGCTGTCAAC  240
Db      193  CTGGGCGAGCTGACGCCAGCCTGCAGACCGGCAAGAGAGCTGCGAGCCTGTACAAC  252
QY      241  ACCGTGGCCACCTGTACTGTGCTGCACGAGAGATCGAGTCCGCGACCAAGAGAGCC  300
Db      253  ACCGTGGCCACCTGTACTGTGCTGCACGAGCCTGCAGCCTGTCAAGAGACCAAGAGAGCC  312
QY      301  CTGGAACAAGATCGAGAGAGAGCAACMACTGCCAGACAAGATCCAGCAGCGGAGCC  360
Db      313  CTGGAAGAAGATCGAGAGAGAGCAACMACTGCCAGAGAAGGCCAGCAGCGCGGCC  372
QY      361  GCGG-----ACAAAGGCAAGGTGAGCCAGAACCTAACCCCATCTGTGAGAACTG  408
Db      373  GCGCGCGGCAACCGCAACAGCAGCAGGTGAGCCAGAACCTAACCCCATCTGTGAGAACTG  432
QY      409  CAGGGCCAGATGTGACACCAAGCCATCAGCCCCCGCACCTCTGAAGCGCTGGTGAAGTG  468
Db      433  CAGGGCCAGATGTGACACCAAGCCATCAGCCCCCGCACCTCTGAAGCGCTGGTGAAGTG  492
QY      469  ATCGAGGAAAGCCTTCAAGCCCGAGGTGATCCCATGTTCAACCGCCTGAGCAGAGGC  528
Db      493  GTGAGAGAGAAAGCCTTCAAGCCCGAGGTGATCCCATGTTCAAGCGCCTGAGCAGAGGC  552
QY      529  GCCACCCCGCGAGACCTGAACACGATGTTGAACACCGTGGCGGCCACCAAGCCCGCATG  588
Db      553  GCCACCCCGCGAGACCTGAACACGATGTTGAACACCGTGGCGGCCACCAAGCCCGCATG  612
QY      589  CAGATGCTGAAGGACACCATCAACGAGAGCGCGCGAGTGGAGACCGGTGACCCCGTG  648
Db      613  CAGATGCTGAAGGAGACCATCAACGAGAGCGCGCGAGTGGAGACCGGTGACCCCGTG  672
QY      649  CAGCGCGCGCCCATCGCCCGCGCGCAGATGCGCGCAGCCCGCGGCAAGCATGCGCGGC  708
Db      673  CAGCGCGCGCCCATCGCCCGCGCGCAGATGCGCGCAGCCCGCGGCAAGCATGCGCGGC  732
QY      709  ACCACGAGACCTGCAAGAGAGATGCGCTGATGACCAAGAACCCCCCATCCCGTG  768
Db      733  ACCACGAGACCTGCAAGAGAGATGCGCTGATGACCAACAAACCCCCCATCCCGTG  792
QY      769  GCGGACATCTAAGCGGTGATCATCTGGGCTGAACAGATCGTGGGATGTACAGC  828
Db      793  GCGGAGATCTAAGCGGTGATCATCTGGGCTGAACAGATCGTGGGATGTACAGC  852
QY      829  CCGGTGAGCATCTGACATCAAGCAGGCGCCCAAGAGAGCCCTTCCGCACTACGTGAC  888
Db      853  CCCACAGCATCTGACATCCGCGCAGGCGCCCAAGAGAGCCCTTCCGCACTACGTGAC  912
QY      889  CGCTTCTTCAAGACCTTGGCGCGCGAGCAGACCAAGAGGTGAAGACTGATGACC  948
Db      913  CGCTTCTACAAGACCTTGGCGCGCTGAGCAGGCGCAAGCAGCAGTGAAGAACTGATGACC  972
QY      949  GACACCTGTGTGCAAGAGCCCAACCCGACTGCAAGACCATCTGCGCGCTCTCGGC  1008
```

Db 973 GAGACCTGCTGTGTCAGAACGCCAACCCCGCACTGCAGAACCATCTCTGAAGGCTCTCGGC 1032
QY 1009 CCGGCGCCGCACTGTGAGAGATGATGATACCGCTTCCAGAGCGCTGGCGGCCCCAGCCAC 1068
Db 1033 CCGCGCGCCACCTTGAGAGATGATGATACCGCTTCCAGAGCGCTGGCGGCCCCAGCCAC 1092
QY 1069 AAGGCCCGCTGTGCTGCGCCGAGCGGATGAGCCAGGCCAACCA-----CCAGCTGATGATG 1122
Db 1093 AAGGCCCGCTGTGCTGCGCCGAGCGGATGAGCCAGAACCCCGGCACTCATGATG 1152
QY 1123 CAGAAAGACAATTCAAGGGCCCCCGCGCATGCTCAAGTGTCTCAACTGCGGCAAGGAG 1182
Db 1153 CAGCGCGCACTTCCGCAACCAAGCGGAAGACCGTCAAGTGTCTCAACTGCGGCAAGGAG 1212
QY 1183 GGGCACAATGCCCCCAACTGCGCGCGCCCCCGCAAGAGGGCTGTGGAAGTGGCGCAAG 1242
Db 1213 GGGCACAACCGCCAGAACTGCGCGCGCCCCCGCAAGAGGGCTGTGCGCGCGC 1272
QY 1243 GAGGGCCACAGATGAAGGACTGCAACCGAGCGCCAGGCCCACTTCTGGGCAAGATCTGG 1302
Db 1273 GAAGGACACCAATGAAGATTGCACTGAGAGACAGGCTAATTTTGAAGAAAGATCTGG 1332
QY 1303 CCCGACCAAGGGCGCGCGCCCACTTCTGCAAGCGCGCCCCGAGCCCAACCGCGCC 1362
Db 1333 CTTCTCAAGAGGAAGGCGCAGGAATTTTCTTCAAGAGACCAAGCCCAACAGCCCA 1392
QY 1363 CCGCGCAGAGCTTCCGCTTGCAGAG-----ACCACCCCGCGCCAGAACAGAGAG--- 1413
Db 1393 CCAGAAAGAGCTTCAAGTTGGGAGAGAAACAACTCCCTCTCAGAACAGAGAGCG 1452
QY 1414 --AGCAAGGACCGCGAGACCTTGACCACTGAAGAGCTGTTCGCAACGACCCCTG 1470
Db 1453 ATAGACAAGAACTGTATCTTAACTTCCCTCAGATCACTCTTGGCAACGACCCCTCG 1512
QY 1471 AGCCAGTAA 1479
Db 1513 TCACAGTAA 1521

RESULT 9
US-09-475-515-79
; Sequence 79, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
; LENGTH: 1865
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GP2
US-09-475-515-79

Query Match 75.4%; Score 1114.6; DB 4; Length 1865;
Best Local Similarity 85.8%; Pred. No. 6.6e-164;
Matches 1295; Conservative 0; Mismatches 184; Indels 30; Gaps 4;

QY 1 ATGGCGCGCCGCGCAGCATCTGCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCGC 60

Db 13 ATGGCGCGCCGCGCAGCGTGTGAGCGCGCGCGAGCTGAGCAAGTGGAGAAAGATCCGC 72
QY 61 CTGCGCGCGCGCAAGAGTGCTCATGATGAAGCACCTGTGTGGCCAGCCCGAG 120
Db 73 CTGCGCGCGCGCAAGAGATGATCAAGCTGAACACATCGTGTGGCCAGCCCGAG 132
QY 121 CTGGAAGAGTTGCGCTTGAACCCCGCTGTGTGAGACCAAGAGGGCTGCAAGCATC 180
Db 133 CTGGAAGCGCTTCCGCGTGAACCCCGCTGTGTGAGACCAAGAGGGCTGCGCGCATC 192
QY 181 ATCCGCGAGCTGCAACCCCGCTGTGAGACCCGCGAGGAGGAGCTGAAGAGCTTTCAAC 240
Db 193 CTGGGCCAGCTGCAAGCCCAAGCTGTGAGACCCGCGAGGAGGAGCTGCGCATCAAC 252
QY 241 ACCGTGGCCACCTGTGTACTGCGTGAACGAGAGATCGAGTCCGCGACACCAAGAGGCC 300
Db 253 ACCGTGGCCACCTGTGTACTGCGTGAACGAGAGATCGAGTCCGCGACCAAGAGGCC 312
QY 301 CTGGAACAAGTTCGAGAGAGGACGAAACAAGTCCAGAGAGATTCAGAGGCGGAGCC 360
Db 313 CTGGAACAAGTTCGAGAGAGGACGAAACAAGTCCAGAGAGGCGGAGGCGCGCC 372
QY 361 GCGC-----ACAAAGGCAAGGTGAGCGCAACTACCCCATCTGTGAGAACTTG 408
Db 373 GCGCGCGCGCACCGGCAACAGCAGCGAGTGAGCCAGAACTACCCCATCTGTGAGAACTTG 432
QY 409 CAGGCGCAGATGTGCAACCAAGGCTATCAGCCCGCGCACCTGAAAGCCTGGGTGAAGGTG 468
Db 433 CAGGCGCAGATGTGCAACCAAGGCTATCAGCCCGCGCACCTGAAAGCCTGGGTGAAGGTG 492
QY 469 ATCAGAGAGAGGCTTCAAGCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGC 528
Db 493 GTGAGAGAGAGGCTTCAAGCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGC 552
QY 529 GCCACCCCGCAGAGCTTGAACACGATGTTGAACACCGTGGCGGCCACAGGCGGCCATG 588
Db 553 GCCACCCCGCAGAGCTTGAACACGATGTTGAACACCGTGGCGGCCACAGGCGGCCATG 612
QY 589 CAGATGCTGAAGGACACCATCAAGAGAGGCGCGAGTGGGACCGCGTGACACCCCGTG 648
Db 613 CAGATGCTGAAGGAGACCATCAAGAGAGGCGCGAGTGGGACCGCGTGACACCCCGTG 672
QY 649 CAGCGCGCGCCATCGCCCCCGCGCAGATGCGCGAGCCCGCGCAGCGACATCGCGGC 708
Db 673 CAGCGCGCGCCATCGCCCCCGCGCAGATGCGCGAGCCCGCGCAGCGACATCGCGGC 732
QY 709 ACCACAGACACCTGCAAGAGAGATCGCTGTGATGACCAAGACACCCCGCATCCCGGTG 768
Db 733 ACCACAGACACCTGCAAGAGAGATCGCTGTGATGACCAAGACACCCCGCATCCCGGTG 792
QY 769 GCGCATCTTAAGCGGTGATCATCTGGGCTTGAACAAGATCGTGCGATGTACAGC 828
Db 793 GCGGAGATCTAAGCGGTGATCATCTGGGCTTGAACAAGATCGTGCGATGTACAGC 852
QY 829 CCGGTAGCATCTGAGACATCAAGAGGCGCCCAAGAGGCTTCCGCGCATCGTGAGC 888
Db 853 CCGACAGCATCTGAGACATCGCGCAGGCGCCCAAGAGGCTTCCGCGCATCGTGAGC 912
QY 889 CGCTTCTTAAGACCTGCGCGCGAGCAGAGACACCAAGAGGTGAAGAACTGATGACC 948
Db 913 CGCTTCTTAAGACCTGCGCGCGAGCAGAGACACCAAGAGGTGAAGAACTGATGACC 972
QY 949 GACACCTGTGTGCAAGACGCAACCCGCACTGCAAGACCATCTGCGCTCTCGGC 1008
Db 973 GAGACCTGTGTGCAAGACGCAACCCGCACTGCAAGACCATCTGAAAGCTCTCGGC 1032
QY 1009 CCGGCGCGCAGCTGAGAGAGATGAGCGCTGCGAGGGCGTGGCGGCCAGCCAC 1068
Db 1033 CCGCGCGCACCTGAGAGAGATGAGCGCTGCGAGGGCGTGGCGGCCAGCCAC 1092
QY 1069 AAGCGCGCGTGTGCGCGAGCGATGAGCGCAACA-----CCAGCGTATGATG 1122

Db 1093 AAGCCCCGCTGCTGGCCGAGGCGATGAGCCAGGTGACGAACCCGGCGACCATCATGATG 1152
QY 1123 CAGAAGAGCACTTCAAGGGCCCCCGCGCATCTGTCAAGTGTTCATCTGGCGAAGAG 1182
Db 1153 CAGCGGGCACTTCCGCAACCAAGCGGAAGACCGTCAAGTGTTCATCTGGCGAAGAG 1212
QY 1183 GGCCACATCGCCCGCAATGCGCGCGCCCCCGCAAGAAAGGGCTGTGGAAGTGGCGCAAG 1242
Db 1213 GGCCACACCGCCGAGAACTGCGCGCGCCCCCGCAAGAAAGGGCTGTGCGCGCGC 1272
QY 1243 GAGGGCCACCAAGATGAAGAACTGCAACCGAGCGCCCACTTCTGGGCAAGATCTGG 1302
Db 1273 GAAGGACACCAATGAAGAACTGTGCACTGAGAGACAGGCTAATTTTAAAGGAAGTCTGG 1332
QY 1303 CCCAGCCACAAGGGCCCCCGGCACTTCTGTGAGAGCGCGCCGAGCCGACCGCCCC 1362
Db 1333 CCTTCTACAAGGGAAGGCCAGGGAATTTTCTTCAGAGCAGACCAAGCCACAGCCCCA 1392
QY 1363 CCCGCCGAGAGCTTCCGCTTCGAGAG-----ACCACCCCCGGCGCAAGCAAGAG--- 1413
Db 1393 CCAGAAGAGAGCTTCAAGTTGGGAGAGAGAAAACAATCCCTCTGAGAAAGCAAGAGCCG 1452
QY 1414 --AGCAAGGACCGCGAGACCCCTGACCAAGCTGAAGAGCCTGTTGGGCAAGCAAGCCCTG 1470
Db 1453 ATAGACAAAGAACTGTATCTTTAACTTCCCTCAGATCACTTTGGCAACGAGCCCTCG 1512
QY 1471 AGCCAGTAA 1479
Db 1513 TCACAGTAA 1521

RESULT 10

US-09-475-515-6
; Sequence 6, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 4319
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: HIV-Gag-polymerase
US-09-475-515-6

Query Match 75.4%; Score 1114.6; DB 4; Length 4319;
Best Local Similarity 85.8%; Pred. No. 6.6e-164;
Matches 1295; Conservative 0; Mismatches 184; Indels 30; Gaps 4;

QY 1 ATGGGCGCCGCGCCAGCATCTGCGCGCGGCAAGCTGAGCGCCTGGAGAGCGCATCCG 60
Db 7 ATGGGCGCCGCGCCAGCGTGTGAGCGCGCGGAGCTGACAAAGTGGAGAAAGATCCGC 66
QY 61 CTGCGCGCGCGCGCAAGAGTGTCTACATGATGAAGCACTGTGTGGCCAGCCGCGAG 120
Db 67 CTGCGCGCGCGCGCAAGAGAGTACAGTGAAGCACTGTGTGGCCAGCCGCGAG 126

QY 121 CTGGAAGAGTTGCCCCCTGAACCCCGGCTGTGTGAGACCAAGGAGGCTGCAAGCATC 180
Db 127 CTGAGAGCGCTTGCCGTGAACCCCGGCTGTGTGAGACCAAGGAGGCTGCGCCAGATC 186
QY 181 ATCCGCCAGCTGACCCCGGCTGCAGACCGGCGAGGAGCTGAAGACCTGTTCAC 240
Db 187 CTGGGCCAGCTGACGCCAGCTGTGACACCGGCGAGGAGCTGCGAGCTGTACAAC 246
QY 241 ACCGTGGCCACCTGTACTGTGCTGACAGAGAAATCGAGTCCGCGACACCAAGAGGCC 300
Db 247 ACCGTGGCCACCTGTACTGTGCTGACACCGCATCGAGCTCAAGACACCAAGAGGCC 306
QY 301 CTGGACAAGATCGAGAGAGAGACAACAAGTCCAGACAGATCCAGAGCGCGAGGCC 360
Db 307 CTGGAAGAAGATGAGAGAGAGACAACAAGTCCAAAGAAGGCCAGAGCGCGCGCC 366
QY 361 GCGG-----ACAAGGCAAGGTGAGCCAGAACTAACCCCATCTGTGAGAACCTG 408
Db 367 GCGGCGGCGCACCGGCAACAGCAGCCAGGTGAGCCAGAACTAACCCCATCTGTGAGAACCTG 426
QY 409 CAGGGCCAGATGATGACACCGGACATCAAGCCCCGCAACCTCTGAACGCTGGGTGAAGGTG 468
Db 427 CAGGGCCAGATGATGACACCGGACATCAAGCCCCGCAACCTCTGAACGCTGGGTGAAGGTG 486
QY 469 ATCGAGGAAGAGCCTTCAAGCCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGGC 528
Db 487 GTGGAAGGAAGAGCCTTCAAGCCCCGAGGTGATCCCATGTTCAGCGCCCTTGAAGGAGGC 546
QY 529 GCCACCCCCCAGAGCCTGAACACGATGTTGAACAACCGTGGGCGGCCACAGGCGGCATG 588
Db 547 GCCACCCCCCAGAGCCTGAACACGATGTTGAACAACCGTGGGCGGCCACAGGCGGCATG 606
QY 589 CAGATGCTGAAGGACACCATCAACGAGAGGCGCGGAGTGGGACCGCGTGCACCCCGTG 648
Db 607 CAGATGCTGAAGGAGACCATCAACGAGAGGCGCGGAGTGGGACCGCGTGCACCCCGTG 666
QY 649 CAGCGCGGCGCCATCGCCCCCGGCGAGTGGCGAGGCCCGCGGCGAGGACATCGCGGC 708
Db 667 CAGCGCGGCGCCATCGCCCCCGGCGAGTGGCGAGGCCCGCGGCGAGGACATCGCGGC 726
QY 709 ACCACAGACCCCTGAGAGAGCATCGCTGTGATGACCAAGACCCCCCATTCGCCGTG 768
Db 727 ACCACAGACCCCTGAGAGAGCATCGCTGTGATGACCAAGACCCCCCATTCGCCGTG 786
QY 769 GCGGACATCTAAGCGGTGATCATCTGGGCTGAACAAGATCGTGGATGTACAGC 828
Db 787 GCGGAGATCTAAGCGGTGATCATCTGGGCTGAACAAGATCGTGGATGTACAGC 846
QY 829 CCGGTGAGCATCTGACATCAAGCGGCGCCCAAGAGCCCTTCGCGCATCGTGAC 888
Db 847 CCCACAGCATCTGACATCGCGCGAGGCGCCCAAGAGCCCTTCGCGCATCGTGAC 906
QY 889 CGCTTCTTAAGACCTGCGCGCGAGCAGACCAACGAGGTGAAGAACTGATGACC 948
Db 907 CGCTTCTTAAGACCTGCGCGCTGAGCAGGCGCAGCAGACGTGAAGAACTGATGACC 966
QY 949 GACACCTGCTGTGAGAAAGCGCAACCCCGACTGCAAGACCATCTGCGGCTCTCGGC 1008
Db 967 GAGACCTGCTGTGAGAAAGCGCAACCCCGACTGCAAGACCATCTGAAAGCTCTCGGC 1026
QY 1009 CCGGCGCGCAGCTGAGAGAGATGATGACCGGCTGCCAGGGGCTGGCGGCCAGCCAC 1068
Db 1027 CCGGCGCGCAGCTGAGAGAGATGATGACCGGCTGCCAGGGGCTGGCGGCCAGCCAC 1086
QY 1069 AAGCGCGCGTGTGGCCGAGGCGATGAGCCAGGCCAACA-----CCAGCGTATGATG 1122
Db 1087 AAGCGCGCGTGTGGCCGAGGCGATGAGCCAGGCGTGAAGAACTGATGATG 1146
QY 1123 CAGAAGAGCACTTCAAGGCGCCCCGCGCATCTGTCAAGTGTTCATCTGCGCAAGAG 1182
Db 1147 CAGCGCGGCACTTCCGCAACCAAGCGGAAGACCGTCAAGTGTTCATCTGCGCAAGAG 1206
QY 1183 GGCCACATCGCCCGCAACTGCGCGCGCCCCCGCAAGAGGCGTGTGAAAGTGGGCAAG 1242


```
Db 1207 GGCCACACCGCCAGGAAGTCCGCGCGCCCCCGCAAGAGGGCTGCTGGCGCTGCGGCCG 1266
QY 1243 GAGGCCACCATGATGAAGACTGCACCGAGCGCCCACTTCTGGGCAAGATCTGG 1302
Db 1267 GAAGGACACCAATGAAAGATTGCACCTGAGAGACAGGCTAATTTTAAAGGAGATCTGG 1326
QY 1303 CCCAGCCACAAGGCGCGCCCCCGCACTTCTGTCAGAGCGCGCCCGCAAGCCACCGCCCC 1362
Db 1327 CTTCTTACAAGGAGGAGCGCCAGGGAATTTTCTTCAAGACAGACCAAGCCACAGCCCCA 1386
QY 1363 CCCGCCGAGAGCTTCCGCTTCGAGAG-----ACCACCCCCGGCCAAGACAGAG--- 1413
Db 1387 CCAGAGAGAGAGCTTCAGGTTTGGGAGAGAGAAAACAATCCCTCTCAGAGCAGAGCCG 1446
QY 1414 ---AGCAAGAGACCGCAGACCCCTGACCAAGCTTGAAGAGCTGTTCCGCAACGAGCCCTG 1470
Db 1447 ATAGACAAAGGAAGTGTATCCTTAACTTCCCTCAGATCACTCTTGGCAACGAGCCCTCG 1506
QY 1471 AGCCAGTAA 1479
Db 1507 TCACAGTAA 1515
```

RESULT 11

```
US-09-475-515-9
; Sequence 9, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1268
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag
; OTHER INFORMATION: common region
US-09-475-515-9
```

```
Query Match 69.3%; Score 1025.2; DB 4; Length 1268;
Best Local Similarity 89.6%; Pred. No. 4.2e-150;
Matches 1131; Conservative 0; Mismatches 113; Indels 18; Gaps 2;
```

```
QY 1 ATGGGCGCCCGCCGACATCTGCGCGCGCGGCAAGCTGGACGCGCTGGAGCGCATCCGC 60
Db 7 ATGGGCGCCCGCCGACGCTGTGAGCGCGCGGCGAGCTGGAACAAGTGGGAAGAGATCCGC 66
QY 61 CTGCGCGCGCGCGCGCAAGAGTGTCTATCATGATGAAGCACTGTGTGTGGGCCAGCGCGAG 120
Db 67 CTGCGCGCGCGCGCGCAAGAGAGTACAGCTGAAGCAATCTGTGTGGGCCAGCGCGAG 126
QY 121 CTGGAAGAAGTTCGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 180
Db 127 CTGGAAGCGCTTCGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 186
QY 181 ATCCGCCAGCTGCACCCCGCGCTGTCAGACCGGCAAGAGAGAGAGAGAGAGAGAGAGAG 240
Db 187 CTGGGCCAGCTGCAGGCCAGCTGTCAGACCGGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 246
```

```
QY 241 ACCGTGGCCACCTGTACTGCGGTGCACGAGAAGATCGAGGTCCGCGACACCAAGAGGCC 300
Db 247 ACCGTGGCCACCTGTACTGCGGTGCACGAGAAGATCGAGGTCCGCGACACCAAGAGGCC 306
QY 301 CTGGAACAAGATCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 307 CTGGAAGAAGATCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 366
QY 361 GCGC-----ACAAGGCAAGGTGAGCCAGAACTACCCCATCTGTGCAAGAACTTG 408
Db 367 GCGCGCGGCAACCGCAACAGCAGCCAGGTGAGCCAGAACTACCCCATCTGTGCAAGAACTTG 426
QY 409 CAGGCGCAGATGTGTCACCAAGGCAATCAGCCCCCGCAACCTTGAAAGCGCTGGTGAAGTG 468
Db 427 CAGGCGCAGATGTGTCACCAAGGCAATCAGCCCCCGCAACCTTGAAAGCGCTGGTGAAGTG 486
QY 469 ATGGAAGAGAGAGGCTTCAAGCCCCGAGGTGATCCCATGTTTCAAGCGCCCTGAGCGAGGGC 528
Db 487 GTGGAAGAGAGAGGCTTCAAGCCCCGAGGTGATCCCATGTTTCAAGCGCCCTGAGCGAGGGC 546
QY 529 GCCACCCCCCAGAGACCTGAACACGATGTTGAACACCGTGGCGGCCACCAAGCGGCCATG 588
Db 547 GCCACCCCCCAGAGACCTGAACACGATGTTGAACACCGTGGCGGCCACCAAGCGGCCATG 606
QY 589 CAGATGCTGAAGACACCATCAAGAGAGCGCGCGAGTGGGACCGCGTGCACCCCGTG 648
Db 607 CAGATGCTGAAGAGACCATCAAGAGAGCGCGCGAGTGGGACCGCGTGCACCCCGTG 666
QY 649 CAGCGCGCGCGCATGCGCCCCCGCGCAGATGCGCGAGGCCCGCGGCAAGCAATCGCGCGC 708
Db 667 CAGCGCGCGCGCATGCGCCCCCGCGCAGATGCGCGAGGCCCGCGGCAAGCAATCGCGCGC 726
QY 709 ACCACAGCAACCTGCAGAGAGAGATGCGCTGATGACCAAGCAACCCCGCATCCCGGTG 768
Db 727 ACCACAGCAACCTGCAGAGAGAGATGCGCTGATGACCAAGCAACCCCGCATCCCGGTG 786
QY 769 GCGGACATCTAAGCGGTGATCATCTGGGCTGAAACAAGATCGTGCAGATGTACAGC 828
Db 787 GCGGAGATCTAAGCGGTGATCATCTGGGCTGAAACAAGATCGTGCAGATGTACAGC 846
QY 829 CCGGTGAGCATCTGACATCAAGCAGGCGCCCAAGAGAGCCCTTCCGCACTACGTGAC 888
Db 847 CCGGTGAGCATCTGACATCAAGCAGGCGCCCAAGAGAGCCCTTCCGCACTACGTGAC 906
QY 889 CGCTTCTTCAAGACCTGCGCGCGCGAGAGACCAAGAGAGTGAAGAACTGATGAC 948
Db 907 CGCTTCTTCAAGACCTGCGCGCGCGAGAGACCAAGAGAGTGAAGAACTGATGAC 966
QY 949 GACACCTGTGTGTGAGAGCGCAACCCCGCATCTGCAAGACCATCTGAAGGCTCTCGGC 1008
Db 967 GACACCTGTGTGTGAGAGCGCAACCCCGCATCTGCAAGACCATCTGAAGGCTCTCGGC 1026
QY 1009 CCGGCGCGCAGCTGAGAGAGATGATGACCGCTGCGAGGCGTGGCGCGCCAGCCAC 1068
Db 1027 CCGGCGCGCAGCTGAGAGAGATGATGACCGCTGCGAGGCGTGGCGCGCCAGCCAC 1086
QY 1069 AAGCGCGCGTGTGCGCGAGCGATGAGCCAGCCAA-----CCAGCGTATGATG 1122
Db 1087 AAGCGCGCGTGTGCGCGAGCGATGAGCCAGCGTGAAGAACCCCGGCAATCATGATG 1146
QY 1123 CAGAAGAGCACTTCAAGGCGCGCGCGCATCTGCAAGTCTTCACTGCGCAAGAG 1182
Db 1147 CAGCGCGGCACTTCCGCAACGAGCGGAAGACCGTCAAGTCTTCACTGCGCAAGAG 1206
QY 1183 GCGCACATGCGCGCAACTGCGCGCGCGCGCGCAAGAGAGGCTGTGAAGTGGCGCAAG 1242
Db 1207 GCGCACACCGCGAAGAACTGCGCGCGCGCGCGCAAGAGGCTGTGCGCGCGCGC 1266
QY 1243 GA 1244
Db 1267 GA 1268
```


RESULT 13
US-09-974-702-1
; Sequence 1, Application US/09974702
; Patent No. 6696291
; GENERAL INFORMATION:
; APPLICANT: Shiver, John W.
; Davies, Mary Ellen
; Freed, Daniel C.
; Liu, Margaret A.
; Perry, Helen C.
; TITLE OF INVENTION: Synthetic HIV Gag Genes
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/974,702
; FILING DATE: 09-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/017,981
; FILING DATE: 03-FEB-1998
; APPLICATION NUMBER: US60/037,854
; FILING DATE: 07-FEB-1997
; APPLICATION NUMBER: GB9705040.5
; FILING DATE: 12-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hand, J. Mark
; REGISTRATION NUMBER: 36,545
; REFERENCE/DOCKET NUMBER: 19730
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-3905
; TELEFAX: 732-594-4720
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1532 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-974-702-1

Query Match 66.0%; Score 976.6; DB 4; Length 1532;
Best Local Similarity 80.2%; Pred. No. 1.4e-142;
Matches 1205; Conservative 0; Mismatches 274; Indels 24; Gaps 4;

Db 190 CTGGGCGAGCTCCAGCCCTCCCTGCAACAGGCTGTGAGGAGTGAGTCCCTGTACAAC 249
QY 241 ACCGTGGCCACCCTGTACTGCTGTGACGAGAAGATCGAGSTCCGCGACACCAAGAGGCC 300
Db 250 ACAGTGGCTACCTGTACTGTGTGTGACACAGAAGATTGATGTGAAGACACCAAGAGGCC 309
QY 301 CTGACAAGATCGAGAGGAGGACGAACAAGTGCCAGCAGAAAGATTCAGAGGCCAGG-- 358
Db 310 CTGAGAAGATTGAGAGGAGGACGAACAAGTCCAGAAGAAGAGGCCAGCAGGCTGTGCT 369
QY 359 -----CCGCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCTGTGACAGAACTTGAGAGGC 414
Db 370 GGCACAGGCAACTCCAGCCAGGTGTCCAGAACTAACCCCATTTGTGACAGAACTTCAGAGGC 429
QY 415 CAGATGTGACCAAGGCCCATCAGCCCCCGCACCTGTGAACGCTTGGGTGAAGGTGATTCAG 474
Db 430 CAGATGTGACCAAGGCCCATCTCCCCCGGACCCTGAATGCTTGGGTGAAGGTGTGAG 489
QY 475 GAGAAAGCCTTCAGCCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGGCGCCACC 534
Db 490 GAGAAAGCCTTCTCCCTGAGGTGATCCCATGTTCTTGCCCTGTGAGGCTGCCACC 549
QY 535 CCCCAGGACCTTGAAACAGATGTTGAACACCGTGGGGCGCCACAGGCGCCATGACAGATG 594
Db 550 CCCCAGGACCTTGAAACAGATGTTGAACACAGTGGGGGGCCATCAGGCTGACATGACAGATG 609
QY 595 CTGAAGGACACCATCAACGAGAGGCGCCGAGTGGGACCGCGTGACACCCCGTGACGCGC 654
Db 610 CTGAAGGAGACCATCAATGAGAGGCTGTGATGGGACAGGCTGATCTGTGACAGCT 669
QY 655 GGCCTCATGCCCCCGCCAGATGCGGAGCGCCGCGGACGACATGCGCGGACAC 714
Db 670 GGCCTCATGCCCCCGCCAGATGAGGAGCCAGGGGCTTGACATTTGCTGGACAC 729
QY 715 AGCACCTGACGAGCAGATGCTGTGATGACCAACACCCCATCCCGTGAGGCGAC 774
Db 730 TCCACCTCCAGAGCAGATGCTGTGATGACCAACACCCCATCCCTGTGGGGGAA 789
QY 775 ATCTACAAGCGGTGATCATCTGGGCTGAAACAAGATCGTGGAGTGTACAGCCCGTG 834
Db 790 ATCTACAAGAGGTGATCATCTGGGCTGAAACAAGATGTGAGATGTACTCCCCACC 849
QY 835 AGCATCTGACATCAAGCAGGCGCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTC 894
Db 850 TCCATCTGACATCAGGACGCGCCCAAGAGCCCTTACGGGACTATGTGACAGTTTC 909
QY 895 TTCGAAGACCTGCGCGCCGAGCAGACACCCAGAGGTGAAGAACTGATGACCGACACC 954
Db 910 TACAAGACCTGAGGGGTGAGCAGGCTCCCAAGAGGTGAAGAACTGATGACAGAGACC 969
QY 955 CTGCTGTGACGAACGCCAACCCGACTGCAAGACCATCTGCGGCTCTCGGCCCGGC 1014
Db 970 CTGCTGTGACGAATGCCAACCTGACTGCAAGACCATCTGAAGGCCCTGGGCCCTGCT 1029
QY 1015 GCCAGCTGAGAGATGATGACCGCTGCGAGGGCGGTGGGGGCCCCAGCCACAAGGCC 1074
Db 1030 GCCACCTGAGAGATGATGACGCTGCGAGGGGGGTGGGGGCCCTGTGACAAAGGCC 1089
QY 1075 CGCGTGTGCGCGAGGCGATGAGCCAGGCCA-----CAAGCTGATGATGACAAG 1128
Db 1090 AGGGTGTGCTGAGGCGCATGTCTCCAGGTGACCAACTCCGCCACATCATGATGACAGG 1149
QY 1129 AGCAACTTCAAGGCGCCCGCGCATGCTCAAGTGTCTTCACTGCGCAAGAGGGCCAC 1188
Db 1150 GGCACCTTCAAGAACAGAGAAAGACAGTGAAGTCTTCACTGTGCAAGGTGGCCAC 1209
QY 1189 ATCGCCCGCAACTGCGCGCGCCCGCAAGAGGGCTGTGAAAGTGCGGCAAGAGGGC 1248
Db 1210 ATTGCCAAGAACTGTAGGGCCCCAGAAAGAGGGCTGTGAAAGTGTGCAAGAGGGC 1269
QY 1249 CACCATGAAGAACTGACCGAGCGCCAGGCCAATTCTGTGGCAAGATCTGGCCAGC 1308

OY		1369 GAGAGCTTCGCTT-----CGAGAGACCACCCCGGCCAAGAAGCAGAG-----AGC	1416
D8		1390 GAGTCCTTAGGTTTGGGAGAGAAAGCACCCCCAGCCAAGACAGAGCCATTGAC	1449
OY		1417 AAGGACCGGAGAACCTTGACCAAGCCTGAAGAGCCTGTTGGCAAACAACCCCCTGAGCCAG	1476
D8		1450 AAGGAGCTGTATCCCCCTGGCCTCCTGAGGTCCCTGTTGGCAAACAACCCCCTCTCCAG	1509
OY		1477 TAA	1479
D8		1510 TAA	1512

```

RESULT 15
US-09-952-060-25
; Sequence 25, Application US/09952060.
; Patent No. 6733993
; GENERAL INFORMATION:
; APPLICANT: Emin, Emilio A.
; APPLICANT: Youil, Rima
; APPLICANT: Bett, Andrew J.
; APPLICANT: Chen, Ling
; APPLICANT: Kaslow, David C.
; APPLICANT: Silver, John W.
; APPLICANT: Toner, Timothy J.
; APPLICANT: Casimiro, Danilo R.
; TITLE OF INVENTION: ENHANCED FIRST GENERATION ADENOVIRUS
; TITLE OF INVENTION: VACCINES EXPRESSING CODON OPTIMIZED HIV1-GAG, POL, NEF AND
; TITLE OF INVENTION: MODIFICATIONS
; FILE REFERENCE: 20747Y
; CURRENT APPLICATION NUMBER: US/09/952,060
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: PCT/US01/28861
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/317,814
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/279,056
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/233,180
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 37474
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA encoding pMRK45 HIV-1 gag, coding
US-09-952-060-25

Query Match          66.0%; Score 976.6; DB 4; Length 37474;
Best Local Similarity 80.2%; Pred. No. 1.4e-142;
Matches 1205; Conservative 0; Mismatches 274; Indels 24; Gaps 4;

Qy      1 ATGGCGCCCGCGCCAGCATCTGCGCGCGCAAGCTGGAACGCTGGAGCGATCCGC 60
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1272 ATGGGTGCTAGGGCTTCTGTGCTGTGTGTGTGAGTGTGACAAAGTGGAGAAATCAGG 1331

Qy      61 CTGGCGCCCGCGCGCGCAAGAAGTGCTACATGATGAAGCACTGGTGTGGCCAGCCGCGAG 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1332 CTGAGGCTGTGTGCGCAAGAAGAATCAAGCTAAAGCAATTGTGTGGGCTCCAGGGAG 1391

Qy      121 CTGGAAGAAGTTCGCTCTGAACCCCGGCTGCTGGAGACCGAGGAGCTGAAGAGCTGTTCAAC 180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1392 CTGGAAGAAGTTCGCTCTGAACCCCGGCTGCTGGAGACCGAGGAGCTGAGAGGAGATC 1451

Qy      181 ATCGCCAGCTGCACCCCGCCCTGACAGCCGGCAGCGAGGAGCTGAAGAGCTGTTCAAC 240
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1452 CTGGCCAGCTCCAGCCCTCCCTGCAACAGGCTCTGAGAGAGCTGAGGTCCCTGTACAAC 1511

Qy      241 ACCGTGGCCACCTGTACTGCGTGACAGAGAAGATCGAGGTCCGCGACCAAGAGAGGCC 300
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db	1512	ACAGTGGCTAACCTGTACTGTGTGCACCAAGAGATTGATGTGAAGGACCAAGAGAGCC	1571
Oy	301	CTGACCAAGATCGAGAGAGACAGAACAAAGTCCAGACAGAAATCCAGACAGCCGAGG--	358
Db	1572	CTGGAGAAAGATTGAGAGAGACAGAACAAATGTCCAAGAAAGAGCCAGCAGGCTGTGCT	1631
Oy	359	----CCGCCGACAAAGGCGAAGGTGAGCCAGAACTACCCCATCGTGCAGAACTGCAGGGC	414
Db	1632	GGCACAGGCAACTCCAGCCAGGTGTCCAGAACTACCCCAATGTGTCAAGAACTTCAGGGC	1691
Oy	415	CAGATGTGCACACGAGCCATCAGCCCCCGCACCTTGAAACGCTTGGGTGAAGGTATCGAG	474
Db	1692	CAGATGTGCACACGAGCCATCTCCCCCGAACCTGAATGCTGGGTGAAGGTGTGAG	1751
Oy	475	GAGAAAGCCTTCAGCCCCGAGGTGATCCCATGTTCAACGCTCAGCCGAGAGGGCGCCACC	534
Db	1752	GAGAAAGCCTTCTCCCTGAGGTGATCCCATGTTCTGTGCCCTGTCTGAGGGTGCACC	1811
Oy	535	CCCCAGGACCTGAACACAGATGTGAACACCGTGGGCGGCCACAGGCGCCATGCAGATG	594
Db	1812	CCCCAGGACCTGAACACAGTCTGAACACAGTGGGGGGCCATCAGCTGCATGCAGATG	1871
Oy	595	CTGAAGGACACCATCAACGAGAGAGGCGCCGAGTGGGACCGCGTGCACCCCGTGCACGCC	654
Db	1872	CTGAAGGAGACCATCAATGAGAGAGGCTGTGAGTGGACAGGCTGCATCTGTGCACGCT	1931
Oy	655	GCCCCCATCGCCCCCGGCGAGATGCGCGAGCCCCGCGCAGCGACATCGCCGGCACCAAC	714
Db	1932	GCCCCCATTTGCCCCCGGCGAGATGAGGAGGCCAGGGGCTCTGACATGTGCTGGCACCAAC	1991
Oy	715	AGCACCCCTGACGAGAGAGAGATCGCTGATGACCAAGAACCCCCCATCCCGTGGGCGAC	774
Db	1992	TCCACCCCTCAAGAGAGAGATTTGGCTGATGACCAACACCCCCCATCCCTGTGGGGGAA	2051
Oy	775	ATCTACAAGCGGTGATCATCTTGGGCTTGAACAAGATCGTGCAGATGTAACGCCCGTG	834
Db	2052	ATCTACAAGAGGTGATCATCTTGGGCTTGAACAAGATTTGAGATGTACTCCCCCACC	2111
Oy	835	AGCATCTTGGACATCAAGCAGAGGGCCCCAAGAGCCCTTCCGACTACGTGACCGCTTC	894
Db	2112	TCCATCTTGACATCAGGCAAGGGCCCCAAGAGCCCTTCAAGGACTATGTGACAGGTTTC	2171
Oy	895	TTCAAGACCCTGCGCGCGCAGACAGACCAAGAGGTGAAGAACTGATGACCGACACC	954
Db	2172	TACAAAGACCCTGAGGGCTGAGCAGGCTCCCAAGAGGTGAAGAACTGATGACAGAGACC	2231
Oy	955	CTGCTGTGTCAGAAACGCCAACCCCGACTGCAAGACCATCTGCGGCTTCGGCCCCGGC	1014
Db	2232	CTGCTGTGTCAGAAATGCCAACCTGTACTGCAAGACCATCTGAAGGCCCTTGGGCCCTGCT	2291
Oy	1015	GCCAGCCTGAGAGAGATGATGACCGCTTGCCAGGGCGTGGGCGGCCCAAGCAAGGCC	1074
Db	2292	GCCACCCCTGAGAGAGATGATGACAGCTTGCCAGGGGCTGGGGGGCCCTGTGCACAAAGCC	2351
Oy	1075	CGCGTGTGCGCGAGGCGATGAGCCAGGCCAAC-----CAGCGTGAATGTCAGAAAG	1128
Db	2352	AGGGTGTGCGCTGAGGCCATGTCTCCAGTGAACCAACTCGGCCACCATCATGATGCAGAGG	2411
Oy	1129	AGCAACTTCAAGGGCCCCCGCGCATCTGTCAAGTGTCTTCACTGCGGCAAGAGGGCCAC	1188
Db	2412	GCGCAACTTCAGGAACCAAGAGGAAGACATGAAGTGTCTTCACTGTGGCAAGTGGGCCAC	2471
Oy	1189	ATCGCCCCGCACTGCGCGCGCCCCCGCAAGAGGGCTGTGGAAGTGCAGCAAGAGGGC	1248
Db	2472	ATTGCCAAGAACTGTAGGGCCCCCGAAGAAAGGGCTGTGGAAGTGTGCAAGAGGGC	2531
Oy	1249	CACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTGGGCAAGATCTGGCCCAGC	1308
Db	2532	CACCAGATGAAGGACTGCATGAAGGCGCCAACTTCTTGGGCAAAATCTGGCCCTTC	2591
Oy	1309	CACAAAGGCGCCCCCGCACTTCTGTGAGAGCCGCCCCGAGCCCCAACCGCCCCCGCC	1368
Db	2592	CACAAAGGCGCGCTGGCACTTCTCTCAATCCAGTCCAGGCTGAGCCCCACAGCCCCCTCCGAG	2651

QY	1369	GAGAGCTTCCGCTT-----CGAGAGACCA	CCCCCGCCAGAA	GAGAG-----AGC	1416
Db	2652	GAGTCCTTCAGT	TTGGGGAGAGAGACCA	CCCCCAGAACGAGAGCCCA	TTGAC 2711
QY	1417	AAGACCGCGAGACCTG	ACCAGCCTGAAGAGCCTG	TTGGCAACGACCCCTGAGCCAG	1476
Db	2712	AAGAGCTGTACCCCTG	GCTCTCCCTGAGGTCCCTG	TTTGGCAACGACCCCTCCTCCAG	2771
QY	1477	TAA	1479		
Db	2772	TAA	2774		

Search completed: September 1, 2005, 22:46:18
Job time : 270.349 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 21:10:07 ; Search time 917.198 Seconds
(without alignments)
10560.254 Million cell updates/sec

Title: US-09-475-704A-3
Perfect score: 1479
Sequence: 1 atggcgccgcgcgcacgcat.....acgacccctgagccagtaa 1479

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 7338684 seqs, 3274456166 residues

Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
- 23: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUBCOMB.seq:*
- 24: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 25: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
- 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1479	100.0	1479	10	US-09-967-464-63 Sequence 63, Appl
2	1479	100.0	1479	10	US-09-899-575-3 Sequence 3, Appl
3	1477.4	99.9	3162	15	US-10-190-435-18 Sequence 18, Appl
4	1477.4	99.9	3462	15	US-10-190-435-16 Sequence 16, Appl
5	1477.4	99.9	4419	15	US-10-190-435-19 Sequence 19, Appl
6	1477.4	99.9	4419	16	US-10-190-305A-14 Sequence 14, Appl
7	1477.4	99.9	4483	15	US-10-190-435-35 Sequence 35, Appl

8	1477.4	99.9	4606	15	US-10-190-435-34 Sequence 34, Appl
9	1477.4	99.9	4615	15	US-10-190-435-36 Sequence 36, Appl
10	1477.4	99.9	4702	15	US-10-190-435-38 Sequence 38, Appl
11	1477.4	99.9	4716	15	US-10-190-435-17 Sequence 17, Appl
12	1477.4	99.9	4716	16	US-10-190-305A-13 Sequence 13, Appl
13	1475.8	99.8	2742	15	US-10-190-435-20 Sequence 20, Appl
14	1475.8	99.8	2742	16	US-10-190-305A-15 Sequence 15, Appl
15	1475.8	99.8	3930	15	US-10-190-435-9 Sequence 9, Appl
16	1475.8	99.8	3930	15	US-10-190-435-10 Sequence 10, Appl
17	1475.8	99.8	3930	15	US-10-190-435-11 Sequence 11, Appl
18	1475.8	99.8	5145	15	US-10-190-435-12 Sequence 12, Appl
19	1475.8	99.8	5145	16	US-10-190-305A-12 Sequence 12, Appl
20	1474.8	99.7	4713	15	US-10-190-435-59 Sequence 59, Appl
21	1474.8	99.7	4713	16	US-10-190-305A-83 Sequence 83, Appl
22	1473.8	99.6	5184	15	US-10-190-435-58 Sequence 58, Appl
23	1473.8	99.6	5184	16	US-10-190-435-82 Sequence 82, Appl
24	1473.4	99.6	1479	10	US-09-967-464-67 Sequence 67, Appl
25	1472.4	99.6	2742	15	US-10-190-435-57 Sequence 57, Appl
26	1472.4	99.6	2742	16	US-10-190-305A-81 Sequence 81, Appl
27	1463	98.9	1479	10	US-09-899-575-20 Sequence 20, Appl
28	1315.2	88.9	3531	15	US-10-190-435-13 Sequence 13, Appl
29	1315.2	88.9	3537	15	US-10-190-435-14 Sequence 14, Appl
30	1315.2	88.9	3537	15	US-10-190-435-15 Sequence 15, Appl
31	1288.8	87.1	1491	10	US-09-899-575-99 Sequence 99, Appl
32	1288	87.1	1494	10	US-09-899-575-51 Sequence 51, Appl
33	1276.8	86.3	1509	10	US-09-899-575-4 Sequence 4, Appl
34	1275.2	86.2	1509	10	US-09-967-464-64 Sequence 64, Appl
35	1273.6	86.1	1509	10	US-09-967-464-68 Sequence 68, Appl
36	1260.8	85.2	1509	10	US-09-899-575-21 Sequence 21, Appl
37	1221.2	82.6	9166	18	US-10-359-120-168 Sequence 168, App
38	1206	81.5	9788	18	US-10-359-120-174 Sequence 174, App
39	1197.8	81.0	1515	17	US-10-387-336-4 Sequence 4, Appl
40	1197.8	81.0	4472	17	US-10-387-336-75 Sequence 75, Appl
41	1197.8	81.0	4608	17	US-10-387-336-76 Sequence 76, Appl
42	1197.8	81.0	4689	17	US-10-387-336-74 Sequence 74, Appl
43	1197.8	81.0	4766	17	US-10-387-336-73 Sequence 73, Appl
44	1196.2	80.9	2799	16	US-10-241-009-18 Sequence 18, Appl
45	1196.2	80.9	2799	16	US-10-190-434B-18 Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-09-967-464-63
; Sequence 63, Application US/09967464
; Publication No. US20030138453A1
; GENERAL INFORMATION:
; APPLICANT: O'Hagan, Derek
; APPLICANT: Otten, Gillis
; APPLICANT: Donnelly, John J.
; APPLICANT: Barnett, Susan
; APPLICANT: Singh, Mamohan
; APPLICANT: Uimer, Jeffrey
; APPLICANT: Dubensky, Jr., Thomas W.
; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS
; FILE REFERENCE: P16269, 004
; CURRENT APPLICATION NUMBER: US/09/967,464
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/236,105
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/315,905
; PRIOR FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 63
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-09-967-464-63
Query Match 100.0%; Score 1479; DB 10; Length 1479;

Best Local Similarity 100.0%; Pred. No. 4.6e-308; Matches 1479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGGCGCCCGCGCGCATCTCTGGCGGGCGGCAAGCTGAGCGCTGGGAGCGCATCCGC 60
Db 1 ATGGGCGCCCGCGCGCATCTCTGGCGGGCGGCAAGCTGAGCGCTGGGAGCGCATCCGC 60
QY 61 CTGGCGCCCGCGCGGCAAGAGTGTCTATGATGAAGCACTGTGTGGGCCAGCGCGAG 120
Db 61 CTGGCGCCCGCGCGGCAAGAGTGTCTATGATGAAGCACTGTGTGGGCCAGCGCGAG 120
QY 121 CTGGAAGAGTTCCGCTGAAACCCCGGCTGTGAGACCAGCGGGCTGCAAGAGATC 180
Db 121 CTGGAAGAGTTCCGCTGAAACCCCGGCTGTGAGACCAGCGGGCTGCAAGAGATC 180
QY 181 ATCGGCCAGCTGCACCCCGCTGAGACCGGCAAGAGAGCTGAAGAGCTGTTC AAC 240
Db 181 ATCGGCCAGCTGCACCCCGCTGAGACCGGCAAGAGAGCTGAAGAGCTGTTC AAC 240
QY 241 ACCGTGGCCACCTGTACTGCGTGAAGAGATGAGTCCGCGACACCAAGAGAGCC 300
Db 241 ACCGTGGCCACCTGTACTGCGTGAAGAGATGAGTCCGCGACACCAAGAGAGCC 300
QY 301 CTGGAACAAGATCGAGAGAGAGACAAGTGCCAGCAGAGATCCAGAGCCGAGGCC 360
Db 301 CTGGAACAAGATCGAGAGAGAGACAAGTGCCAGCAGAGATCCAGAGCCGAGGCC 360
QY 361 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCTGTGAGAACCTGAGGGCCAGATG 420
Db 361 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCTGTGAGAACCTGAGGGCCAGATG 420
QY 421 GTGACACGAGGCCATCAGCCCCCGCACCTGAACGCTGGGTGAAGGTGATGAGAGAA 480
Db 421 GTGACACGAGGCCATCAGCCCCCGCACCTGAACGCTGGGTGAAGGTGATGAGAGAA 480
QY 481 GCCTTCAGCCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGGCCACCCCCAG 540
Db 481 GCCTTCAGCCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGGCCACCCCCAG 540
QY 541 GACCTGAACACGATGTGAACACCGTGGGGCGGCAAGCGCCCATGAGATGCTGAAG 600
Db 541 GACCTGAACACGATGTGAACACCGTGGGGCGGCAAGCGCCCATGAGATGCTGAAG 600
QY 601 GACACCATCAACGAGAGAGCGCGCGAGTGGGACCGCGTGACCCGCTGACGCGGCCCC 660
Db 601 GACACCATCAACGAGAGAGCGCGCGAGTGGGACCGCGTGACCCGCTGACGCGGCCCC 660
QY 661 ATGCCCCCGGCGCAGATGCGGAGCGCGCGGCAAGCATCGCGGCAACCAAGCAACC 720
Db 661 ATGCCCCCGGCGCAGATGCGGAGCGCGCGGCAAGCATCGCGGCAACCAAGCAACC 720
QY 721 CTGCAAGAGCAGATCGCTGATGACCAAGCAACCCCATCCCGTGGGGCAATCTTAC 780
Db 721 CTGCAAGAGCAGATCGCTGATGACCAAGCAACCCCATCCCGTGGGGCAATCTTAC 780
QY 781 AAGCGGTGATCATCTGGGCTGAAACAAGATCGTGGGATGTACAGCCCGTGAACATC 840
Db 781 AAGCGGTGATCATCTGGGCTGAAACAAGATCGTGGGATGTACAGCCCGTGAACATC 840
QY 841 CTGACATCAAGCAGAGCGCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAG 900
Db 841 CTGACATCAAGCAGAGCGCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAG 900
QY 901 ACCCTGGCGCGCGAGAGACCAAGAGGTGAAGAACTGATGACCGACACCTGTGCTG 960
Db 901 ACCCTGGCGCGCGAGAGACCAAGAGGTGAAGAACTGATGACCGACACCTGTGCTG 960
QY 961 GTGAGAAAGCCCAACCCCGACTGCAAGACCATCTGTGGCGCTCTGGCCCCGGCGCAGC 1020
Db 961 GTGAGAAAGCCCAACCCCGACTGCAAGACCATCTGTGGCGCTCTGGCCCCGGCGCAGC 1020
QY 1021 CTGAGAGAGATGATACCGCTGCAAGGGCGTGGCGGCCCAAGCAAGGCCCGCTG 1080
Db 1021 CTGAGAGAGATGATACCGCTGCAAGGGCGTGGCGGCCCAAGCAAGGCCCGCTG 1080
```

```
Db 1021 CTGAGAGAGATGATACCGCTGCAAGGGCGGTGGGGCGGCCCAAGCAAGGCCCGCTG 1080
QY 1081 CTGGCCGAGCGGATGAGCCAGGCCAACACACCGCTGATGATGCAAGAGCACTTCAAG 1140
Db 1081 CTGGCCGAGCGGATGAGCCAGGCCAACACACCGCTGATGATGCAAGAGCACTTCAAG 1140
QY 1141 GGCCCCCGGCGCATCGTCAAGTGTCTTCACTGCGGCAAGAGAGGCCACATCGCCGCAAC 1200
Db 1141 GGCCCCCGGCGCATCGTCAAGTGTCTTCACTGCGGCAAGAGAGGCCACATCGCCGCAAC 1200
QY 1201 TGCCCGGCCCCCGCAAGAGGGCTGTGAAAGTGCCGCAAGAGAGGCCACAGATGAAG 1260
Db 1201 TGCCCGGCCCCCGCAAGAGGGCTGTGAAAGTGCCGCAAGAGAGGCCACAGATGAAG 1260
QY 1261 GACTGCACCGAGCGCCAGGCCAATTCTGTGGCAAGATGTGGCCCAAGCAAGGGCCGC 1320
Db 1261 GACTGCACCGAGCGCCAGGCCAATTCTGTGGCAAGATGTGGCCCAAGCAAGGGCCGC 1320
QY 1321 CCGGCAACTTCTGCAAGAGCGCGCCGAGCCACCGCCCCCGCGGAGCTTCCGC 1380
Db 1321 CCGGCAACTTCTGCAAGAGCGCGCCGAGCCACCGCCCCCGCGGAGCTTCCGC 1380
QY 1381 TTCGAGGAGACCAACCCCGGCAAGAGAGAGCAAGAGACCGCGAGACCTTGAACAGC 1440
Db 1381 TTCGAGGAGACCAACCCCGGCAAGAGAGAGCAAGAGACCGCGAGACCTTGAACAGC 1440
QY 1441 CTGAAGAGCTGTTCGGCAACGACCCCTGAGCCAGTAA 1479
Db 1441 CTGAAGAGCTGTTCGGCAACGACCCCTGAGCCAGTAA 1479
```

RESULT 2
US-09-899-575-3
; Sequence 3, Application US/09899575
; Publication No. US20030223961A1

; GENERAL INFORMATION:
; APPLICANT: Zur Megele, Jan
; APPLICANT: Barnett, Susan W.
; APPLICANT: Egnelbrecht, Susan
; APPLICANT: van Rensburg, Estrelita Janse
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; FILE REFERENCE: PP01631.102
; CURRENT APPLICATION NUMBER: US/09/899,575
; CURRENT FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 09/475,704
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag
US-09-899-575-3

Query Match 100.0%; Score 1479; DB 10; Length 1479;
Best Local Similarity 100.0%; Pred. No. 4.6e-308; Matches 1479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGGCGCCCGCGCGCATCTCTGGCGGGCGGCAAGCTGAGCGCTGGGAGCGCATCCGC 60
Db 1 ATGGGCGCCCGCGCGCATCTCTGGCGGGCGGCAAGCTGAGCGCTGGGAGCGCATCCGC 60
QY 61 CTGGCGCCCGCGCGGCAAGAGTGTCTACATGATGAAGCAACCTGTGTGGGCCAGCGCGAG 120
Db 61 CTGGCGCCCGCGCGGCAAGAGTGTCTACATGATGAAGCAACCTGTGTGGGCCAGCGCGAG 120
QY 121 CTGGAAGAGTTCCGCTGAAACCCCGGCTGTGAGACCAGCGAGGGCTGCAAGAGATC 180
Db 121 CTGGAAGAGTTCCGCTGAAACCCCGGCTGTGAGACCAGCGAGGGCTGCAAGAGATC 180
```



```
QY 181 ATCCGCCAGCTGCACCCCGCCCTGCGACCGGACGAGAGCTGAAGCCTGTTCAAC 240
Db 181 ATCCGCCAGCTGCACCCCGCCCTGCGACCGGACGAGAGCTGAAGCCTGTTCAAC 240
QY 241 ACCGTGGCCACCTGTACTGCTGTGCGACGAGAAGATCGAGTCCGGCACCAAGAGGCC 300
Db 241 ACCGTGGCCACCTGTACTGCTGTGCGACGAGAAGATCGAGTCCGGCACCAAGAGGCC 300
QY 301 CTGGAAGAATCGAGAGAGAGAGAGCAAGTGCAGAGAGAGATCCAGAGGCCGAGGCC 360
Db 301 CTGGAAGAATCGAGAGAGAGAGAGCAAGTGCAGAGAGAGATCCAGAGGCCGAGGCC 360
QY 361 GCCGCAAGGGGCAAGGTGAGCCAGAACTAACCCATCGTGCAGAACTGCAGGGCCAGATG 420
Db 361 GCCGCAAGGGGCAAGGTGAGCCAGAACTAACCCATCGTGCAGAACTGCAGGGCCAGATG 420
QY 421 GTGCACCAAGGCCATCAACCCCGCCACCTGAACGCTGGGTGAAGGTGATCGAGAGAA 480
Db 421 GTGCACCAAGGCCATCAACCCCGCCACCTGAACGCTGGGTGAAGGTGATCGAGAGAA 480
QY 481 GCCTTACGCCCCCGAGGTGATCCCACTGTTCACCGCCCTGAGCGAGGGCCCAACCCCCAG 540
Db 481 GCCTTACGCCCCCGAGGTGATCCCACTGTTCACCGCCCTGAGCGAGGGCCCAACCCCCAG 540
QY 541 GACCTGAACACGATGTTGAACACCGGTGGGGGGCCACCAAGGCCCATGCAATGCTGAAG 600
Db 541 GACCTGAACACGATGTTGAACACCGGTGGGGGGCCACCAAGGCCCATGCAATGCTGAAG 600
QY 601 GACACCATCAACGAGAGAGAGGGCCCGCGAGTGGGACCGCGTGCACCCCGTGCACGCCGCC 660
Db 601 GACACCATCAACGAGAGAGAGGGCCCGCGAGTGGGACCGCGTGCACCCCGTGCACGCCGCC 660
QY 661 ATCGCCCCCGCCAGATGCGCGAGCGCCCGGCGAGGACATCGCCGGCAACCAAGCAACC 720
Db 661 ATCGCCCCCGCCAGATGCGCGAGCGCCCGGCGAGGACATCGCCGGCAACCAAGCAACC 720
QY 721 CTGCAAGAGACGATGCGCTTGATGACGACCAACCCCTCATCCCGTGGGGCGACATCTAC 780
Db 721 CTGCAAGAGACGATGCGCTTGATGACGACCAACCCCTCATCCCGTGGGGCGACATCTAC 780
QY 781 AAGCGGTGATCATCTGGGCTTGAACAAGATCGTGGGATGTACAGCCCGTGAGCATC 840
Db 781 AAGCGGTGATCATCTGGGCTTGAACAAGATCGTGGGATGTACAGCCCGTGAGCATC 840
QY 841 CTGACATCAAGCAGGCCCCCAAGGACCTTCCGCACTACGTGACCGCTTCTTCAAG 900
Db 841 CTGACATCAAGCAGGCCCCCAAGGACCTTCCGCACTACGTGACCGCTTCTTCAAG 900
QY 901 ACCGTGGCGCCGAGCAGACCCCAAGAGGTGAAGAACTGATGACCAACCTGCTG 960
Db 901 ACCGTGGCGCCGAGCAGACCCCAAGAGGTGAAGAACTGATGACCAACCTGCTG 960
QY 961 GTGCAGAACGCCAACCCGACTGCAAGACCATCTGCGGCTCTCGGCCCGCGCGCAGC 1020
Db 961 GTGCAGAACGCCAACCCGACTGCAAGACCATCTGCGGCTCTCGGCCCGCGCGCAGC 1020
QY 1021 CTGAGAGATGATGACCGCTGCGAGGCGTGGCGGCCCGCAGCCAAGGCCCGGTG 1080
Db 1021 CTGAGAGATGATGACCGCTGCGAGGCGTGGCGGCCCGCAGCCAAGGCCCGGTG 1080
QY 1081 CTGCGCAGGCGATGAGCCAGGCCAACACAGCGTGTATGACAGAGAACAACTTCAAG 1140
Db 1081 CTGCGCAGGCGATGAGCCAGGCCAACACAGCGTGTATGACAGAGAACAACTTCAAG 1140
QY 1141 GGCCCCCGCGCATCTCAAGTGTCTCACTGCGCAAGAGGGCCCATCGCCCGCAAC 1200
Db 1141 GGCCCCCGCGCATCTCAAGTGTCTCACTGCGCAAGAGGGCCCATCGCCCGCAAC 1200
QY 1201 TGCCGCGCCCCCGCAAGAGGGCTGCTGAAGTGCAGGCAAGAGGCCACAGATGAAG 1260
Db 1201 TGCCGCGCCCCCGCAAGAGGGCTGCTGAAGTGCAGGCAAGAGGCCACAGATGAAG 1260
```

```
QY 1261 GACTGCACCGAGCGCCAGGCCAACTTCTTGGGCAAGATCTGGGCCAGCCACAAAGGCCGC 1320
Db 1261 GACTGCACCGAGCGCCAGGCCAACTTCTTGGGCAAGATCTGGGCCAGCCACAAAGGCCGC 1320
QY 1321 CCGGCAACTTCTTGCAGAGAGCGCCCGAGAGCCACCGCCCCCGCCGAGAGCTTCCGC 1380
Db 1321 CCGGCAACTTCTTGCAGAGAGCGCCCGAGAGCCACCGCCCCCGCCGAGAGCTTCCGC 1380
QY 1381 TTCGAGGAGACCAACCCCGGCCAGAAAGCAGAGAGCAAGAGACCGCGAGACCTGAACAGC 1440
Db 1381 TTCGAGGAGACCAACCCCGGCCAGAAAGCAGAGAGCAAGAGACCGCGAGACCTGAACAGC 1440
QY 1441 CTGAAGAGCCTGTTGCGCAACGACCCCTGAGCCAGTAA 1479
Db 1441 CTGAAGAGCCTGTTGCGCAACGACCCCTGAGCCAGTAA 1479
```

```
RESULT 3
US-10-190-435-18
; Sequence 18, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEDEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; FILE REFERENCE: P18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 3162
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagRTmut_C
US-10-190-435-18
```

```
Query Match 99.9%; Score 1477.4; DB 15; Length 3162;
Best Local Similarity 99.9%; Pred. No. 9.7e-308;
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGGCGCCCGCCAGCATCTCGCGCGCGCGCAAGCTGACGCTGGAGCGATCCGC 60
Db 7 ATGGGGCGCCCGCCAGCATCTCGCGCGCGCGCAAGCTGACGCTGGAGCGATCCGC 66
QY 61 CTGGCGCCCGCGCGCAAGAGTGTACATGATGAAGCACCTGTGTGGCCAGCCGAG 120
Db 67 CTGGCGCCCGCGCGCAAGAGTGTACATGATGAAGCACCTGTGTGGCCAGCCGAG 126
QY 121 CTGGAAGAATTCGCCCTGAACCCCGGCTGCTGGAAGACCAAGGAGGCTGCAAGCATC 180
Db 127 CTGGAAGAATTCGCCCTGAACCCCGGCTGCTGGAAGACCAAGGAGGCTGCAAGCATC 186
QY 181 ATCCGCCAGCTGACCCCGCTGCAAGACCGGCGAGCGAGAGCTGAAGAGCTGTTCAAC 240
Db 187 ATCCGCCAGCTGACCCCGCTGCAAGACCGGCGAGCGAGAGCTGAAGAGCTGTTCAAC 246
QY 241 ACCGTGGCCACCTGTACTGCGTGCAGAGAGATCGAGTCCGGCAGACCAAGAGGCC 300
Db 247 ACCGTGGCCACCTGTACTGCGTGCAGAGAGATCGAGTCCGGCAGACCAAGAGGCC 306
QY 301 CTGGAACAAGTGAAGAGAGAGCAACAAGTCCAGCAGAGATCCAGAGGCCGAGGCC 360
Db 307 CTGGAACAAGTGAAGAGAGAGCAACAAGTCCAGCAGAGATCCAGAGGCCGAGGCC 366
QY 361 GCCGCAAGGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACTGCAAGGCCAGATG 420
Db 367 GCCGCAAGGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACTGCAAGGCCAGATG 426
```

```
QY 421 GTGACCAAGGCCATCAGCCCCCGCACCCCTGAAACGCTGGGTGAAGGTGATCGAGAGAAG 480
Db 427 GTGACACCAAGGCCATCAGCCCCCGCACCCCTGAAACGCTGGGTGAAGGTGATCGAGAGAAG 486
QY 481 GCCTTCAGCCCCGAGGTGATCCCATGTTCAACGGCCCTGAGCGAGGGCGCCACCCCCCAG 540
Db 487 GCCTTCAGCCCCGAGGTGATCCCATGTTCAACGGCCCTGAGCGAGGGCGCCACCCCCCAG 546
QY 541 GACCTGAACACGATGTTGAACACCGTGGCGGGCCACCAAGGCCCGCATGCAATGCTGAAG 600
Db 547 GACCTGAACACGATGTTGAACACCGTGGCGGGCCACCAAGGCCCGCATGCAATGCTGAAG 606
QY 601 GACACCATCAACGAGGAGGCGCGCGAGTGGGACCGCGTGCAACCCCGTGCAACCGCGGCC 660
Db 607 GACACCATCAACGAGGAGGCGCGCGAGTGGGACCGCGTGCAACCCCGTGCAACCGCGGCC 666
QY 661 ATGCCCCCGGCGCAGATGCGCGAGCGCGCGGCGAGCGACATGCGCCGGCAACCAAGACA 720
Db 667 ATGCCCCCGGCGCAGATGCGCGAGCGCGCGGCGAGCGACATGCGCCGGCAACCAAGACA 726
QY 721 CTGACGAGCAGATCGCTGGATGACCAACAACCCCATCCCGTGGGCGACATCTAC 780
Db 727 CTGACGAGCAGATCGCTGGATGACCAACAACCCCATCCCGTGGGCGACATCTAC 786
QY 781 AAGCGGTGATCATCTGGGGCTGAAACAAGATCGTGGGATGTACAGCCCCGTGAGCATC 840
Db 787 AAGCGGTGATCATCTGGGGCTGAAACAAGATCGTGGGATGTACAGCCCCGTGAGCATC 846
QY 841 CTGACATCAAGCAGGCGCCCAAGGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAG 900
Db 847 CTGACATCAAGCAGGCGCCCAAGGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAG 906
QY 901 ACCCTGCGCGCGAGCAGACACCCAGAGGTGAAGAACTGGATGACGACACCTGCTG 960
Db 907 ACCCTGCGCGCGAGCAGACACCCAGAGGTGAAGAACTGGATGACGACACCTGCTG 966
QY 961 GTGCAAGACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTGCGCCCGCGCGCAGC 1020
Db 967 GTGCAAGACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTGCGCCCGCGCGCAGC 1026
QY 1021 CTGAGAGAGATGATGACCGCTGCGCAGGCGGTGGCGGCCAGGCAAAAGCCCGGTG 1080
Db 1027 CTGAGAGAGATGATGACCGCTGCGCAGGCGGTGGCGGCCAGGCAAAAGCCCGGTG 1086
QY 1081 CTGCGCGAGGCGATGAGCCAGGCCAACACAGCGTGTATGACAGAAGAACCTTCAAG 1140
Db 1087 CTGCGCGAGGCGATGAGCCAGGCCAACACAGCGTGTATGACAGAAGAACCTTCAAG 1146
QY 1141 GGCCCCCGGCGCATCTCAAGTGTCTTCAACTGCGGCAAGAGGGCCACATGCCCCGCAAC 1200
Db 1147 GGCCCCCGGCGCATCTCAAGTGTCTTCAACTGCGGCAAGAGGGCCACATGCCCCGCAAC 1206
QY 1201 TGCCGCGCCCCCGCAAGAGGCTGTGGAAGTGCAGCAAGAGGGCCACCAAGATGAAG 1260
Db 1207 TGCCGCGCCCCCGCAAGAGGCTGTGGAAGTGCAGCAAGAGGGCCACCAAGATGAAG 1266
QY 1261 GACTGCAACGAGCGCGCAAGCCACTTCTGGGCAAGATCTGCCCCACCAAGGGCGCG 1320
Db 1267 GACTGCAACGAGCGCGCAAGCCACTTCTGGGCAAGATCTGCCCCACCAAGGGCGCG 1326
QY 1321 CCGGCAACTTCTCTGAGAGCGCGCGCGAGCCCAACCGCCCCCGCGGAGAGCTTCCG 1380
Db 1327 CCGGCAACTTCTCTGAGAGCGCGCGCGAGCCCAACCGCCCCCGCGGAGAGCTTCCG 1386
QY 1381 TTGAGAGAGACCAACCCCGGCGCAGAACAGAGACAGAACCGCGGAGACCTTGACCAAG 1440
Db 1387 TTGAGAGAGACCAACCCCGGCGCAGAACAGAGACAGAACCGCGGAGACCTTGACCAAG 1446
QY 1441 CTGAAGAGCCTGTTGGCAACGACCCCTGAGCCAGTAA 1479
Db 1447 CTGAAGAGCCTGTTGGCAACGACCCCTGAGCCAGAAA 1485
```

```
RESULT 4
US-10-190-435-16
; Sequence 16, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEHDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; FILE REFERENCE: P18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 3462
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagProteinRtmu_C
US-10-190-435-16
```

```
Query Match 99.9%; Score 1477.4; DB 15; Length 3462;
Best Local Similarity 99.9%; Pred. No. 9,7e-308;
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ATGGGCGCGCGCGCGCAGCATCTGCGCGCGCGCAAGCTGAGCGCTGGAGCGATCCGC 60
Db 7 ATGGGCGCGCGCGCGCAGCATCTGCGCGCGCGCAAGCTGAGCGCTGGAGCGATCCGC 66
QY 61 CTGCGCGCGCGCGCGCAAGAGTCTACATGATGAAGCACTGTGTGGGCCAGCCGCGAG 120
Db 67 CTGCGCGCGCGCGCGCAAGAGTCTACATGATGAAGCACTGTGTGGGCCAGCCGCGAG 126
QY 121 CTGAGAAAGTTGCGCCCTGAACCCCGGCTGTGGAAGACCAAGGCGTGAAGCATC 180
Db 127 CTGAGAAAGTTGCGCCCTGAACCCCGGCTGTGGAAGACCAAGGCGTGAAGCATC 186
QY 181 ATCCGCCAGCTGACACCCCGCTGACAGACCGGCAAGAGAGCTGAAGAGCTGTTCAAC 240
Db 187 ATCCGCCAGCTGACACCCCGCTGACAGACCGGCAAGAGAGCTGAAGAGCTGTTCAAC 246
QY 241 ACCGTGGCCACCCTGTACTGTGCTGACAGAGAATCGAGGTCCGACACCAAGAGAGCC 300
Db 247 ACCGTGGCCACCCTGTACTGTGCTGACAGAGAATCGAGGTCCGACACCAAGAGAGCC 306
QY 301 CTGGAACAAGATGAGAGGAGGAGAAACAAGTGCACAGAGAATCCAGAGGCCGAGGCC 360
Db 307 CTGGAACAAGATGAGAGGAGGAGAAACAAGTGCACAGAGAATCCAGAGGCCGAGGCC 366
QY 361 GCGGACAAGGGCAAGGTGAGCCAGAACTACCCCATCTGTGCAAACTGACAGGCCAGATG 420
Db 367 GCGGACAAGGGCAAGGTGAGCCAGAACTACCCCATCTGTGCAAACTGACAGGCCAGATG 426
QY 421 GTGACCAAGGCCATCAGCCCGGCAACCTGAAAGCGCTGGTGAAGGTGATGAGAGAAAG 480
Db 427 GTGACCAAGGCCATCAGCCCGGCAACCTGAAAGCGCTGGTGAAGGTGATGAGAGAAAG 486
QY 481 GCCTTCAGCCCGGAGGTGATCCCATGTTTCAACCGCCCTGAGCGAGGCGCCACCCCCAG 540
Db 487 GCCTTCAGCCCGGAGGTGATCCCATGTTTCAACCGCCCTGAGCGAGGCGCCACCCCCAG 546
QY 541 GACCTGAACACGATGTTGAACACCGTGGCGGGCCACCAAGGCCCGCATGCAATGCTGAAG 600
Db 547 GACCTGAACACGATGTTGAACACCGTGGCGGGCCACCAAGGCCCGCATGCAATGCTGAAG 606
QY 601 GACACCATCAACGAGGAGGCGCGCGAGTGGGACCGGTGCAACCCCGTGACCGCGGCCCC 660
Db 607 GACACCATCAACGAGGAGGCGCGCGAGTGGGACCGGTGCAACCCCGTGACCGCGGCCCC 666
```

```
QY 661 ATCGCCCCCGCCAGATGCGCGGAGCCCCCGCGGAGCATCGCCGGCACCAGCACC 720
    |||
Db 667 ATCGCCCCCGCCAGATGCGCGGAGCCCCCGCGGAGCATCGCCGGCACCAGCACC 726
QY 721 CTGAGAGAGAGATCGCCTTGATGATGACCAACCCCCCATCCCCCGTGGGCGACATCTAC 780
    |||
Db 727 CTGAGAGAGAGATCGCCTTGATGATGACCAACCCCCCATCCCCCGTGGGCGACATCTAC 786
QY 781 AAGCGGTGATCATCTGGGCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 840
    |||
Db 787 AAGCGGTGATCATCTGGGCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 846
QY 841 CTGACATCAAGCAGGCCCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAG 900
    |||
Db 847 CTGACATCAAGCAGGCCCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAG 906
QY 901 ACCCTGCGCGCGGAGCAGCACCAGAGAGTGAAGAACTGATGACCGACACCTGCTG 960
    |||
Db 907 ACCCTGCGCGCGGAGCAGCACCAGAGAGTGAAGAACTGATGACCGACACCTGCTG 966
QY 961 GTGCAAGACGCCAACCCCGACTGCAAGAACCATCTCGCGCGCTCTCGGCCCCGGCGCAGC 1020
    |||
Db 967 GTGCAAGACGCCAACCCCGACTGCAAGAACCATCTCGCGCGCTCTCGGCCCCGGCGCAGC 1026
QY 1021 CTGAGAGAGATGATGATCGCCTGCGCAGGCGGTGGGCGGCCCCAGCCACAAGGCCGCTG 1080
    |||
Db 1027 CTGAGAGAGATGATGATCGCCTGCGCAGGCGGTGGGCGGCCCCAGCCACAAGGCCGCTG 1086
QY 1081 CTGCGCGAGGCGGATGAGCCAGGCCAACCAACCGGTGATGATGACAGAGCAACTTCAAG 1140
    |||
Db 1087 CTGCGCGAGGCGGATGAGCCAGGCCAACCAACCGGTGATGATGACAGAGCAACTTCAAG 1146
QY 1141 GGGCCCCGGCGCATGTCAAGTCTTCAACTGCGGCAAGAGGCGCCCATGCGCCGCAAC 1200
    |||
Db 1147 GGGCCCCGGCGCATGTCAAGTCTTCAACTGCGGCAAGAGGCGCCCATGCGCCGCAAC 1206
QY 1201 TGCGCGCGCCCCCGCAAGAGGGCTGTGGAAGTGCGGCAAGGAGGCCACCAAGTGAAG 1260
    |||
Db 1207 TGCGCGCGCCCCCGCAAGAGGGCTGTGGAAGTGCGGCAAGGAGGCCACCAAGTGAAG 1266
QY 1261 GACTGCAACCGAGCGCGCAAGGCTTCTGCGCAAGATCTGCGCCAGCCACCAAGGCGCGC 1320
    |||
Db 1267 GACTGCAACCGAGCGCGCAAGGCTTCTGCGCAAGATCTGCGCCAGCCACCAAGGCGCGC 1326
QY 1321 CCGGCAACTTCTCTGAGAGCGCGCGCGAGCCACCGCCCCCGCGAGAGCTTCCGC 1380
    |||
Db 1327 CCGGCAACTTCTCTGAGAGCGCGCGCGAGCCACCGCCCCCGCGAGAGCTTCCGC 1386
QY 1381 TTGAGAGAGACCAACCCCGCGCAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
    |||
Db 1387 TTGAGAGAGACCAACCCCGCGCAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1446
QY 1441 CTGAAGAGCCTGTTGCGCAACGACCCCTGAGCCAGTAA 1479
    |||
Db 1447 CTGAAGAGCCTGTTGCGCAACGACCCCTGAGCCAGAAA 1485
```

```
RESULT 5
US-10-190-435-19
; Sequence 19, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MESEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES, AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
```

```
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 4419
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagRTmutatRevNef_C
US-10-190-435-19

Query Match      99.9%; Score 1477.4; DB 15; Length 4419;
Best Local Similarity 99.9%; Pred. No. 9.6e-308;
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGCGCGCGCCAGCATCCTGCGCGCGCGCAAGCTGAGCCTGGAGCGCATCCGC 60
    |||
Db 7 ATGGCGCGCGCGCCAGCATCCTGCGCGCGCGCAAGCTGAGCCTGGAGCGCATCCGC 66
QY 61 CTGCGCGCGCGCGCAAGAGTGTACATGATGAAGCACCTGCTGGGCGCAGCGCGAG 120
    |||
Db 67 CTGCGCGCGCGCGCAAGAGTGTGTACATGATGAAGCACCTGCTGGGCGCAGCGCGAG 126
QY 121 CTGAGAGAGTTCGCTTGAACCCCGGCTGCTGAGAGACGAGCGGCTGCAAGCATC 180
    |||
Db 127 CTGAGAGAGTTCGCTTGAACCCCGGCTGCTGAGAGACGAGCGGCTGCAAGCATC 186
QY 181 ATCCGCCAGCTGACACCCCGCTGACAGACCGGCGAGCGAGAGCTGAAGAGCTGTTCAAC 240
    |||
Db 187 ATCCGCCAGCTGACACCCCGCTGACAGACCGGCGAGCGAGAGCTGAAGAGCTGTTCAAC 246
QY 241 ACCGTGGCCACCCCTGTACTGCTGTCACGAGAGAGATCGAGTCCGCGACACCAAGAGGCC 300
    |||
Db 247 ACCGTGGCCACCCCTGTACTGCTGTCACGAGAGAGATCGAGTCCGCGACACCAAGAGGCC 306
QY 301 CTGAGCAAGATTCAGAGAGAGAGCAACAAGTCCAGCAGAGATCCAGAGCGCGAGGCC 360
    |||
Db 307 CTGAGCAAGATTCAGAGAGAGAGCAACAAGTCCAGCAGAGATCCAGAGCGCGAGGCC 366
QY 361 GCCGCAAGAGGCAAGGTGAGGCCAAGTACCCCATCTGTCAGAACTGCAAGGCCAGATG 420
    |||
Db 367 GCCGCAAGAGGCAAGGTGAGGCCAAGTACCCCATCTGTCAGAACTGCAAGGCCAGATG 426
QY 421 GTGCAACCAAGCCATCAGCCCCCGCACCTGTAAACGCTGTGTAAGGTGATCGAGAGAA 480
    |||
Db 427 GTGCAACCAAGCCATCAGCCCCCGCACCTGTAAACGCTGTGTAAGGTGATCGAGAGAA 486
QY 481 GCCTTCAGCCCCGAGGTGATCCCCCATGTTCAACCGCCCTGAGCGAGGGCGCCACCCCCAG 540
    |||
Db 487 GCCTTCAGCCCCGAGGTGATCCCCCATGTTCAACCGCCCTGAGCGAGGGCGCCACCCCCAG 546
QY 541 GACCTGAACACGATGTTGAACAACCGTGGGCGGCCACAGGCGCCATGACAGTGTGAAG 600
    |||
Db 547 GACCTGAACACGATGTTGAACAACCGTGGGCGGCCACAGGCGCCATGACAGTGTGAAG 606
QY 601 GACACCATCAACGAGAGAGCGCGCGAGTGGGACCGCGTGACCCCGTGACGCGCGCCCC 660
    |||
Db 607 GACACCATCAACGAGAGAGCGCGCGAGTGGGACCGCGTGACCCCGTGACGCGCGCCCC 666
QY 661 ATCGCCCCCGCGCAGATGCGGAGCCCCCGCGCAGGACATCGCCGGCACCAGCACC 720
    |||
Db 667 ATCGCCCCCGCGCAGATGCGGAGCCCCCGCGCAGGACATCGCCGGCACCAGCACC 726
QY 721 CTGAGAGAGAGATCGCCTTGATGATGACCAACCCCCCATCCCCCGTGGGCGACATCTAC 780
    |||
Db 727 CTGAGAGAGAGATCGCCTTGATGATGACCAACCCCCCATCCCCCGTGGGCGACATCTAC 786
QY 781 AAGCGGTGATCATCTGGGCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 840
    |||
Db 787 AAGCGGTGATCATCTGGGCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATC 846
QY 841 CTGACATCAAGCAGGCCCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAG 900
    |||
Db 847 CTGACATCAAGCAGGCCCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAG 906
```

QY 901 ACCCTGCGCGCCGAGAGACACCCAGAGGTGAAGAACTGGATGACCGACACCTGCTG 960
DB 907 ACCCTGCGCGCCGAGAGACACCCAGAGGTGAAGAACTGGATGACCGACACCTGCTG 966
QY 961 GTGCAGAAACGCCAACCCCGACTGCAAGAACCATCTCTGCGCGCTCTGCGCCCCGCGCGCAGC 1020
DB 967 GTGCAGAAACGCCAACCCCGACTGCAAGAACCATCTCTGCGCGCTCTGCGCCCCGCGCGCAGC 1026
QY 1021 CTGAGAGAGATGATGACCGCCTGCGAGGGCGTGGGGCGGCCCAAGGCCCGCGTG 1080
DB 1027 CTGAGAGAGATGATGACCGCCTGCGAGGGCGTGGGGCGGCCCAAGGCCCGCGTG 1086
QY 1081 CTGCGCGAGGCGATGAGCCAGGCCAACACAGCGGTGATGACAGAGAGCAACTTCAAG 1140
DB 1087 CTGCGCGAGGCGATGAGCCAGGCCAACACAGCGGTGATGACAGAGAGCAACTTCAAG 1146
QY 1141 GGCCCCCGCGCATCTCAAGTCTTCAACTGCGCAAGAGGCCCATGCGCCGCAAC 1200
DB 1147 GGCCCCCGCGCATCTCAAGTCTTCAACTGCGCAAGAGGCCCATGCGCCGCAAC 1206
QY 1201 TGCCCGCGCCCCCGCAAGAGGGCTGTGAAGTGGCGCAAGAGGGCCACCAAGTGAAG 1260
DB 1207 TGCCCGCGCCCCCGCAAGAGGGCTGTGAAGTGGCGCAAGAGGGCCACCAAGTGAAG 1266
QY 1261 GACTGCACCGAGCGCGCCCAACTTCTTGGGCAAGATCTGGCCCAAGAGGGCCGC 1320
DB 1267 GACTGCACCGAGCGCGCCCAACTTCTTGGGCAAGATCTGGCCCAAGAGGGCCGC 1326
QY 1321 CCGGCACTTCTCTGAGAGCGCGCCGAGACCCACCGCCCCCGCGAGAGCTTCCGC 1380
DB 1327 CCGGCACTTCTCTGAGAGCGCGCCGAGACCCACCGCCCCCGCGAGAGCTTCCGC 1386
QY 1381 TTGAGAGAGACCAACCCCGCGCCAGAGACAGAGAGCAAGACCGCGAGACCTGACAGC 1440
DB 1387 TTGAGAGAGACCAACCCCGCGCCAGAGACAGAGAGCAAGACCGCGAGACCTGACAGC 1446
QY 1441 CTGAAGAGCTGTTCGGCAACGACCCCTTGAGCCAGTAA 1479
DB 1447 CTGAAGAGCTGTTCGGCAACGACCCCTTGAGCCAGAAA 1485

RESULT 6
US-10-190-305A-14
; Sequence 14, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190,305A
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 4419
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: GagRTmutatRevNef_C
US-10-190-305A-14

Query Match 99.9%; Score 1477.4; DB 16; Length 4419;
Best Local Similarity 99.9%; Pred. No. 9.6e-308;
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGCCGCGCCAGCATCTGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCGC 60
DB 7 ATGGGCGCCGCGCCAGCATCTGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCGC 66

QY 61 CTGCGCCCCGCGCAAGAACTGCTACATGATGAAGCACTGTGTGGGCCAGCCGCGAG 120
DB 67 CTGCGCCCCGCGCAAGAACTGCTACATGATGAAGCACTGTGTGGGCCAGCCGCGAG 126
QY 121 CTGAGAAAGTTCCCTCTGAACCCCGGCTGTGTGAGACCAAGAGGGCTGCAAGCAGATC 180
DB 127 CTGAGAAAGTTCCCTCTGAACCCCGGCTGTGTGAGACCAAGAGGGCTGCAAGCAGATC 186
QY 181 ATCCGCCAGCTGCACCCCGCCTGCAGACCGCGCAGCGAGAGCTGAAGAGCTGTTCAAC 240
DB 187 ATCCGCCAGCTGCACCCCGCCTGCAGACCGCGCAGCGAGAGCTGAAGAGCTGTTCAAC 246
QY 241 ACCGTGGCCACCTGTACTGTGCTGCACGAGAGATGAGAGTCCGACACCAAGAGAGCC 300
DB 247 ACCGTGGCCACCTGTACTGTGCTGCACGAGAGATGAGAGTCCGACACCAAGAGAGCC 306
QY 301 CTGGAACAAGATCGAGAGGAGGAGCAACAAGTGCAGAGAAATCCAGAGAGCCGAGGCC 360
DB 307 CTGGAACAAGATCGAGAGGAGGAGCAACAAGTGCAGAGAAATCCAGAGAGCCGAGGCC 366
QY 361 GCCGACAAGGGCAAGGTGAGGCCAAGAACTACCCCATCTGTGCAGAACTGCAAGGCCAGATG 420
DB 367 GCCGACAAGGGCAAGGTGAGGCCAAGAACTACCCCATCTGTGCAGAACTGCAAGGCCAGATG 426
QY 421 GTGCACCAAGGCCATCAGCCCCCGCACCCCTGAACGCTGTGTAAGGTGATCGAGAGAAAG 480
DB 427 GTGCACCAAGGCCATCAGCCCCCGCACCCCTGAACGCTGTGTAAGGTGATCGAGAGAAAG 486
QY 481 GCCTTCAGCCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCAG 540
DB 487 GCCTTCAGCCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCAG 546
QY 541 GACCTGAACACGATGTTGAACAAGTGGGGGCCAACAGCGGCCCATGCAATGCTGAAG 600
DB 547 GACCTGAACACGATGTTGAACAAGTGGGGGCCAACAGCGGCCCATGCAATGCTGAAG 606
QY 601 GACACCATCAACGAGAGAGGCGCGAGTGGGACCGCGTGCAACCCCGTGACGCGCGCCCC 660
DB 607 GACACCATCAACGAGAGAGGCGCGAGTGGGACCGCGTGCAACCCCGTGACGCGCGCCCC 666
QY 661 ATGCCCCCGGCAAGATGCGGAGCCCCCGGCGCAGCAATCGCCGCAACCAAGCAGCAGC 720
DB 667 ATGCCCCCGGCAAGATGCGGAGCCCCCGGCGCAGCAATCGCCGCAACCAAGCAGCAGC 726
QY 721 CTGCAGAGACAGATCGCCTGTGATGACCAACCCCCCATCCCGTGCGCAGCATCTAC 780
DB 727 CTGCAGAGACAGATCGCCTGTGATGACCAACCCCCCATCCCGTGCGCAGCATCTAC 786
QY 781 AAGCGGTGATCATCTGGGCTGAACAAGATCGTGCAGATGTACAGCCCGTGAGCATC 840
DB 787 AAGCGGTGATCATCTGGGCTGAACAAGATCGTGCAGATGTACAGCCCGTGAGCATC 846
QY 841 CTGGAATCAAGAGAGGGCCCCCAAGAGGCCCTTCGCGCACTACGTGAGACCGCTTTCAAG 900
DB 847 CTGGAATCAAGAGAGGGCCCCCAAGAGGCCCTTCGCGCACTACGTGAGACCGCTTTCAAG 906
QY 901 ACCCTGCGCGCGCAGACAGACCCCAAGAGGTGAAGAACTGGATGACCGACACCTGCTG 960
DB 907 ACCCTGCGCGCGCAGACAGACCCCAAGAGGTGAAGAACTGGATGACCGACACCTGCTG 966
QY 961 GTGCAGAAACGCCAACCCCGACTGCAAGAACCATCTCTGCGCGCTCTGCGCCCCGCGCGCAGC 1020
DB 967 GTGCAGAAACGCCAACCCCGACTGCAAGAACCATCTCTGCGCGCTCTGCGCCCCGCGCGCAGC 1026
QY 1021 CTGAGAGAGATGATGACCGCCTGCGAGGGCGTGGGGCGGCCCAAGGCCCGCGTG 1080
DB 1027 CTGAGAGAGATGATGACCGCCTGCGAGGGCGTGGGGCGGCCCAAGGCCCGCGTG 1086
QY 1081 CTGCGCGAGGCGATGAGCCAGGCCAACACAGCGGTGATGACAGAGAGCAACTTCAAG 1140
DB 1087 CTGCGCGAGGCGATGAGCCAGGCCAACACAGCGGTGATGACAGAGAGCAACTTCAAG 1146

QY 1141 GGGCCCCCGCGCATCGTCAAGTGTCTCACTGCGGCAAGAGGGCCACATGCCCCGCAAC 1200
|
|
|
Db 1147 GGGCCCCCGCGCATCGTCAAGTGTCTCACTGCGGCAAGAGGGCCACATGCCCCGCAAC 1206
|
|
|
QY 1201 TGCCCGCCCCCGCGCAAGAGGGCTGTGGAAGTGCAGGCAAGAGGGCCACCATGTAAG 1260
|
|
|
Db 1207 TGCCCGCCCCCGCGCAAGAGGGCTGTGGAAGTGCAGGCAAGAGGGCCACCATGTAAG 1266
|
|
|
QY 1261 GACTGCACCGAGCGGCCCAACTTCTTGAGCAAGATCTGGCCCAAGGCCCAAGGGCCGC 1320
|
|
|
Db 1267 GACTGCACCGAGCGGCCCAACTTCTTGAGCAAGATCTGGCCCAAGGCCCAAGGGCCGC 1326
|
|
|
QY 1321 CCGGCAACTTCTCTGAGAGCGCGCCGAGCCCAACCGCCCCCGCGGAGAGCTTCCGC 1380
|
|
|
Db 1327 CCGGCAACTTCTCTGAGAGCGCGCCGAGCCCAACCGCCCCCGCGGAGAGCTTCCGC 1386
|
|
|
QY 1381 TTGAGAGAGACCAACCCCGGCGCAGAAAGAGAGAGAGACCGCGAGACCTGACAGC 1440
|
|
|
Db 1387 TTGAGAGAGACCAACCCCGGCGCAGAAAGAGAGAGAGACCGCGAGACCTGACAGC 1446
|
|
|
QY 1441 CTGAAGAGCTGTTCGGCAACGACCCCTGAGCCAGTAA 1479
|
|
|
Db 1447 CTGAAGAGCTGTTCGGCAACGACCCCTGAGCCAGAAA 1485
|
|
|

RESULT 7

US-10-190-435-35
; Sequence 35, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEDEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35
; LENGTH: 4483
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160mod.TV1.dv1V2-gagmod.BM965
US-10-190-435-35

Query Match 99.9%; Score 1477.4; DB 15; Length 4483;
Best Local Similarity 99.9%; Pred. No. 9.6e-308;
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGCCCGCGCGCATCTGCGCGCGCGCGCAAGCTGACGCTGGAGCGCATCCGC 60
|
|
|
Db 3005 ATGGCGCCCGCGCGCATCTGCGCGCGCGCGCAAGCTGACGCTGGAGCGCATCCGC 3064
|
|
|
QY 61 CTGCGCCCGCGCGCGCAAGAGTGTATCATGATGAAGCACTGTGTGGCCAGCGCGAG 120
|
|
|
Db 3065 CTGCGCCCGCGCGCGCAAGAGTGTATCATGATGAAGCACTGTGTGGCCAGCGCGAG 3124
|
|
|
QY 121 CTGAGAAGTTGCGCTGAACCCCGGCTGTGAGACCAAGGAGCTGCAAGCAGATC 180
|
|
|
Db 3125 CTGAGAAGTTGCGCTGAACCCCGGCTGTGAGACCAAGGAGCTGCAAGCAGATC 3184
|
|
|
QY 181 ATCCGCGAGCTGACCCCGCTGAGACCGGCGAGGAGCTGAAGAGCTGTTCAAC 240
|
|
|
Db 3185 ATCCGCGAGCTGACCCCGCTGAGACCGGCGAGGAGCTGAAGAGCTGTTCAAC 3244
|
|
|
QY 241 ACCGTGGCCACCTGTACTGCGTGCAAGAGATGAGGTCCGCGACACCAAGAGGCGC 300
|
|
|

Db 3245 ACCGTGGCCACCTGTACTGCGTGCAAGAGATGAGGTCCGCGACACCAAGAGGCGC 3304
|
|
|
QY 301 CTGGAACAAGATGAGAGAGAGAGCAACAAGTGTCCAGAGAGATCCAGAGCGCGAGGCC 360
|
|
|
Db 3305 CTGGAACAAGATGAGAGAGAGAGCAACAAGTGTCCAGAGAGATCCAGAGCGCGAGGCC 3364
|
|
|
QY 361 GCCGCAAGGGCAAGGTGAGCCAGAACTACCCATCGTGACAGAACCTGACGGCCAGATG 420
|
|
|
Db 3365 GCCGCAAGGGCAAGGTGAGCCAGAACTACCCATCGTGACAGAACCTGACGGCCAGATG 3424
|
|
|
QY 421 GTGACCAAGGCCATCAGCCCCCGCACCTGAAAGCCTGGGTGAAGGTGATGAGAGAAAG 480
|
|
|
Db 3425 GTGACCAAGGCCATCAGCCCCCGCACCTGAAAGCCTGGGTGAAGGTGATGAGAGAAAG 3484
|
|
|
QY 481 GCCTTCAGCCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCAG 540
|
|
|
Db 3485 GCCTTCAGCCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCAG 3544
|
|
|
QY 541 GACCTGAACAGATGTTGAACAACGTGGCGGCCACAGGCGCCATGACAGATGCTGAAG 600
|
|
|
Db 3545 GACCTGAACAGATGTTGAACAACGTGGCGGCCACAGGCGCCATGACAGATGCTGAAG 3604
|
|
|
QY 601 GACACCATCAACAGAGAGAGCGCGCGAGTGGGACCGCGTGCAACCCCGTGACGCGCGCCC 660
|
|
|
Db 3605 GACACCATCAACAGAGAGAGCGCGCGAGTGGGACCGCGTGCAACCCCGTGACGCGCGCCC 3664
|
|
|
QY 661 ATCGCCCCCGCGCAGATGCGCGAGCCCGCGCGCAGGACATCGCCGCGCACCAAGCAGACC 720
|
|
|
Db 3665 ATCGCCCCCGCGCAGATGCGCGAGCCCGCGCGCAGGACATCGCCGCGCACCAAGCAGACC 3724
|
|
|
QY 721 CTGAGAGAGCAGATCGCTGATGACCAAGCAACCCCCCATCCCGTGCGGCGCATCTAC 780
|
|
|
Db 3725 CTGAGAGAGCAGATCGCTGATGACCAAGCAACCCCCCATCCCGTGCGGCGCATCTAC 3784
|
|
|
QY 781 AAGCGGTGATCATCTGGGCTGAACAAGATCGTGCGGATGACAGCCCGTGAAGCATC 840
|
|
|
Db 3785 AAGCGGTGATCATCTGGGCTGAACAAGATCGTGCGGATGACAGCCCGTGAAGCATC 3844
|
|
|
QY 841 CTGACATCAAGCAGAGGCGCCCAAGAGGCCCTCCGCGACTACGTGAGACCGCTTCAAG 900
|
|
|
Db 3845 CTGACATCAAGCAGAGGCGCCCAAGAGGCCCTCCGCGACTACGTGAGACCGCTTCAAG 3904
|
|
|
QY 901 ACCCTGCGCGCGCAGAGCAGACCCAGAGGTGAAGACTGATGACCGCACCCCTGCTG 960
|
|
|
Db 3905 ACCCTGCGCGCGCAGAGCAGACCCAGAGGTGAAGACTGATGACCGCACCCCTGCTG 3964
|
|
|
QY 961 GTGCAGAACCGCAACCCCGACTGCAAGCAACCTCTGCGGCTCTCGGCCCCGCGCCAGC 1020
|
|
|
Db 3965 GTGCAGAACCGCAACCCCGACTGCAAGCAACCTCTGCGGCTCTCGGCCCCGCGCCAGC 4024
|
|
|
QY 1021 CTGAGAAGATGATGACCGCTGCGCAGGCGGTGGGCGGCCCCAGCCACAAGGCGCGTG 1080
|
|
|
Db 4025 CTGAGAAGATGATGACCGCTGCGCAGGCGGTGGGCGGCCCCAGCCACAAGGCGCGTG 4084
|
|
|
QY 1081 CTGCGCCGAGGCGATGAGCCAGGCCAACAACGCGTGATGATGACAGAGGCAACTTCAAG 1140
|
|
|
Db 4085 CTGCGCCGAGGCGATGAGCCAGGCCAACAACGCGTGATGATGACAGAGGCAACTTCAAG 4144
|
|
|
QY 1141 GGGCCCCCGGCGCATGTCAAGTCTTCAACTGCGGCAAGAGGCGCCACATGCCCCGCAAC 1200
|
|
|
Db 4145 GGGCCCCCGGCGCATGTCAAGTCTTCAACTGCGGCAAGAGGCGCCACATGCCCCGCAAC 4204
|
|
|
QY 1201 TGCCCGCGCCCCCGCAAGAGGGCTGTGGAAGTGCAGGCAAGAGGGCCACCATGTAAG 1260
|
|
|
Db 4205 TGCCCGCGCCCCCGCAAGAGGGCTGTGGAAGTGCAGGCAAGAGGGCCACCATGTAAG 4264
|
|
|
QY 1261 GACTGCACCGAGCGCCAGGCCAATTCTTGAGCAAGATCTGGCCCAAGCCCAAGGGCGCGC 1320
|
|
|
Db 4265 GACTGCACCGAGCGCCAGGCCAATTCTTGAGCAAGATCTGGCCCAAGCCCAAGGGCGCGC 4324
|
|
|
QY 1321 CCGGCAACTTCTCTGAGAGCGCGCCGAGGCCACCGCCCCCGCGCGAGAGCTTCCGC 1380
|
|
|
Db 4325 CCGGCAACTTCTCTGAGAGCGCGCCGAGGCCACCGCCCCCGCGCGAGAGCTTCCGC 4384
|
|
|

OY 1381 TTCGAGAGACCAACCCCGCCAGAGCAGAGAGACCGCGAGACCCCTGACCAGC 1440
|
DB 4385 TTCGAGAGACCAACCCCGCCAGAGCAGAGAGACCGCGAGACCCCTGACCAGC 4444
OY 1441 CTGAAGACCTGTTCGGCAACGACCCCTGAGCCAGTAA 1479
|
DB 4445 CTGAAGACCTGTTCGGCAACGACCCCTGAGCCAGTAA 4483

RESULT 8

US-10-190-435-34

; Sequence 34, Application US/10190435

; Publication No. US20030143248A1

; GENERAL INFORMATION:

; APPLICANT: ZUR MEDEDE, Jan

; APPLICANT: BARNETT, Susan W.

; APPLICANT: LIAN, Ying

; APPLICANT: ENGELBRECHT, Susan

; APPLICANT: VAN RENSBURG, Estrelita J.

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C

; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF

; FILE REFERENCE: PP18133.003 / 2302-18133

; CURRENT APPLICATION NUMBER: US/10/190,435

; NUMBER OF SEQ ID NOS: 319

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 34

; LENGTH: 4606

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:

; OTHER INFORMATION: gp160mod.TV1.dvl-gagmod.BW965

US-10-190-435-34

; OTHER INFORMATION: gp160mod.TV1.dvl-gagmod.BW965

US-10-190-435-34

Query Match 99.9%; Score 1477.4; DB 15; Length 4606;
Best Local Similarity 99.9%; Pred. No. 9.6e-308;
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ATGGGCGCCCGCGCCAGCATCTGCGGGGGCGGCAAGCTGGACGCGCTGGAGCGCATCCGC 60
|
DB 3128 ATGGGCGCCCGCGCCAGCATCTGCGGGGGCGGCAAGCTGGACGCGCTGGAGCGCATCCGC 3187
OY 61 CTGCGCCCGCGCGGCAAGAGTGCTATCATGATGAAGCACTGTGTGGGCCAGCCGCGAG 120
|
DB 3188 CTGCGCCCGCGCGGCAAGAGTGCTATCATGATGAAGCACTGTGTGGGCCAGCCGCGAG 3247
OY 121 CTGAGAAAGTTCCGCTGAAACCCCGGCTGCTGAGACCAAGGCGCTGCAAGCATC 180
|
DB 3248 CTGAGAAAGTTCCGCTGAAACCCCGGCTGCTGAGACCAAGGCGCTGCAAGCATC 3307
OY 181 ATCCGCGAGCTGACCCCGCCCTGAGACCGGCGAGGAGCTGAAGACCTGTTCAAC 240
|
DB 3308 ATCCGCGAGCTGACCCCGCCCTGAGACCGGCGAGGAGCTGAAGACCTGTTCAAC 3367
OY 241 ACCGTGGCCACCTGTATGCTGCGTCAAGAAAGTCAAGTCCGCGACCAAGAGAGCC 300
|
DB 3368 ACCGTGGCCACCTGTATGCTGCGTCAAGAAAGTCAAGTCCGCGACCAAGAGAGCC 3427
OY 301 CTGACAAAGATCGAGAGAGCAGAACTAAGTCCAGCAGAAAGATCCAGCAGGCGGAGCC 360
|
DB 3428 CTGACAAAGATCGAGAGAGCAGAACTAAGTCCAGCAGAAAGATCCAGCAGGCGGAGCC 3487
OY 361 GCCGACAAAGGCGAAGTGAGCCAGAACTAAGTCCAGCAGAACTGCAAGGCGGAGATG 420
|
DB 3488 GCCGACAAAGGCGAAGTGAGCCAGAACTAAGTCCAGCAGAACTGCAAGGCGGAGATG 3547
OY 421 GTGACACGAGCCATCAGCCCGCCAGCCTGAACGCTGGGTGAAGGTGATCGAGAGAAG 480
|
DB 3548 GTGACACGAGCCATCAGCCCGCCAGCCTGAACGCTGGGTGAAGGTGATCGAGAGAAG 3607
OY 481 GCCTTCAGCCCGAGGTGATCCCATGTTCAACGCGCTGAGCGAGGCGCCACCCCCAG 540

DB 3608 GCCTTCAGCCCGAGGTGATCCCATGTTCAACGCGCTGAGCGAGGCGCCACCCCCAG 3667
OY 541 GACCTGAACAGATGTTGAACACCGTGGGGGCCACAGGCGCCATGCAATGCTGAAG 600
|
DB 3668 GACCTGAACAGATGTTGAACACCGTGGGGGCCACAGGCGCCATGCAATGCTGAAG 3727
OY 601 GACACCATCAAGAGAGGCGCGAGTGGACCGCGTGACCCCGTGACCGCGGCC 660
|
DB 3728 GACACCATCAAGAGAGGCGCGAGTGGACCGCGTGACCCCGTGACCGCGGCC 3787
OY 661 ATCGCCCGCGCGAGATGCGCGAGCCCGCGGACGACATCGCGGACCAAGCAGCACC 720
|
DB 3788 ATCGCCCGCGCGAGATGCGCGAGCCCGCGGACGACATCGCGGACCAAGCAGCACC 3847
OY 721 CTGACAGAGCAGATCGCTGATGACCAACCCCGCATCCCGTGGCGCATCTAC 780
|
DB 3848 CTGACAGAGCAGATCGCTGATGACCAACCCCGCATCCCGTGGCGCATCTAC 3907
OY 781 AAGCGTGATCATCTGGGCTGAACAAGATCGTGGATGTACAGCCCGTGAGCATC 840
|
DB 3908 AAGCGTGATCATCTGGGCTGAACAAGATCGTGGATGTACAGCCCGTGAGCATC 3967
OY 841 CTGACATCAAGCAGAGGCGCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTTCAAG 900
|
DB 3968 CTGACATCAAGCAGAGGCGCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTTCAAG 4027
OY 901 ACCCTGCGCGCGAGCAGAGCAACCGAGGTGAAGAACTGATGACCGACACCTGCTG 960
|
DB 4028 ACCCTGCGCGCGAGCAGAGCAACCGAGGTGAAGAACTGATGACCGACACCTGCTG 4087
OY 961 GTGACAGACGCCAACCCCGACTGCAAGACCACTCTGCGGCTCTGCGCCCGCGCGCAGC 1020
|
DB 4088 GTGACAGACGCCAACCCCGACTGCAAGACCACTCTGCGGCTCTGCGCCCGCGCGCAGC 4147
OY 1021 CTGAGAGATGATGACCGCTGCGAGGCGGTGGCGGCCCGACCAAGGCGCGGTG 1080
|
DB 4148 CTGAGAGATGATGACCGCTGCGAGGCGGTGGCGGCCCGACCAAGGCGCGGTG 4207
OY 1081 CTGCGCGAGCGCATGAGCCAGGCCAACACACAGCTGATGATGCAAGAGCACTTCAAG 1140
|
DB 4208 CTGCGCGAGCGCATGAGCCAGGCCAACACACAGCTGATGATGCAAGAGCACTTCAAG 4267
OY 1141 GGCCCCCGCGCATGCTCAAGTCTTCAACTGCGGCAAGAGGCGCACATGCGCCGCAAC 1200
|
DB 4268 GGCCCCCGCGCATGCTCAAGTCTTCAACTGCGGCAAGAGGCGCACATGCGCCGCAAC 4327
OY 1201 TGCCGCGCCCGCCGCAAGAGGCGCTGGAAGTGCGGCAAGAGGCGCACAGATGAAG 1260
|
DB 4328 TGCCGCGCCCGCCGCAAGAGGCGCTGGAAGTGCGGCAAGAGGCGCACAGATGAAG 4387
OY 1261 GACTGCACCGAGCGCCAGGCCAATTCTGCGCAAGATCTGCGCCAGCAAGAGGCGCGC 1320
|
DB 4388 GACTGCACCGAGCGCCAGGCCAATTCTGCGCAAGATCTGCGCCAGCAAGAGGCGCGC 4447
OY 1321 CCGGCAACTTCTGCAAGCGCGCCGAGCCACCGCGCGCGCGCGCGAGCTTCCGC 1380
|
DB 4448 CCGGCAACTTCTGCAAGCGCGCGCGAGCCACCGCGCGCGCGCGCGAGCTTCCGC 4507
OY 1381 TTCGAGAGACCAACCCCGCGCAAGAGCAGAGAGCAAGACCGCGAGACCTGACAGC 1440
|
DB 4508 TTCGAGAGACCAACCCCGCGCAAGAGCAGAGAGCAAGACCGCGAGACCTGACAGC 4567
OY 1441 CTGAAGAGCCTGTTCCGCAACGACCCCTGAGCCAGTAA 1479
|
DB 4568 CTGAAGAGCCTGTTCCGCAACGACCCCTGAGCCAGTAA 4606

RESULT 9

US-10-190-435-36

; Sequence 36, Application US/10190435

; Publication No. US20030143248A1

; GENERAL INFORMATION:

```
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYPEPTIDES ENCODING ANTIGENIC HIV TYPE C
; FILE REFERENCE: P18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36
; LENGTH: 4615
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160mod.TV1.dv2-gagmod.BW965
; US-10-190-435-36
```

Query Match 99.9%; Score 1477.4; DB 15; Length 4615;

Best Local Similarity 99.9%; Pred. No. 9.6e-308;

Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGGGCGCGCGCGCCAGCATCTGCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCG 60
Db 3137 ATGGGCGCGCGCGCCAGCATCTGCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCG 3196
QY 61 CTGCGCGCGCGCGCGCAAGAGTGTCTACTATGATGAAGCACTGTGTGGGCCAGCGCGAG 120
Db 3197 CTGCGCGCGCGCGCGCAAGAGTGTCTACTATGATGAAGCACTGTGTGGGCCAGCGCGAG 3256
QY 121 CTGAGAGAAGTTGCGCCTTGAAACCCCGGCTGTGAGACCAAGGAGGCTGCAAGCAGATC 180
Db 3257 CTGAGAGAAGTTGCGCCTTGAAACCCCGGCTGTGAGACCAAGGAGGCTGCAAGCAGATC 3316
QY 181 ATCCGCGAGCTGCACCCCGCCCTGCAGACCGGCGAGGAGCTGAAGAGCTGTTCAAC 240
Db 3317 ATCCGCGAGCTGCACCCCGCCCTGCAGACCGGCGAGGAGCTGAAGAGCTGTTCAAC 3376
QY 241 ACCGTGGCCACCTGTACTGCGTGCAAGAAAGATCGAGGTCCGCGACCAAGAGAGCC 300
Db 3377 ACCGTGGCCACCTGTACTGCGTGCAAGAAAGATCGAGGTCCGCGACCAAGAGAGCC 3436
QY 301 CTGGAACAAGATCGAGAGGAGCAGAACAAGTGCCAGCAGAAAGATCCAGCGCGCAGGCC 360
Db 3437 CTGGAACAAGATCGAGAGGAGCAGAACAAGTGCCAGCAGAAAGATCCAGCGCGCAGGCC 3496
QY 361 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCTGCAGAACTGCAGGGCCAGATG 420
Db 3497 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCTGCAGAACTGCAGGGCCAGATG 3556
QY 421 GTGACACAGGCGCATCAGCCCCCGCACCTGAACGCTGGGTGAAGGTGATCGAGAGAAG 480
Db 3557 GTGACACAGGCGCATCAGCCCCCGCACCTGAACGCTGGGTGAAGGTGATCGAGAGAAG 3616
QY 481 GCCTTCAGCCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGCGCCACCCCCAG 540
Db 3617 GCCTTCAGCCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGCGCCACCCCCAG 3676
QY 541 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAAGCGCCCATGAGATGCTGAAG 600
Db 3677 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAAGCGCCCATGAGATGCTGAAG 3736
QY 601 GACACCATCAACGAGAGGCGCGCGAGTGGAGCCGCTGCACCCCGTGCAGCGCGGCC 660
Db 3737 GACACCATCAACGAGAGGCGCGCGAGTGGAGCCGCTGCACCCCGTGCAGCGCGGCC 3796
QY 661 ATGCGCGCGCGCGCAGATGCGCGAGCCCGCGGAGCGACATCGCGCGCACCAACGAGACC 720
Db 3797 ATGCGCGCGCGCGCAGATGCGCGAGCCCGCGGAGCGACATCGCGCGCACCAACGAGACC 3856
```

```
QY 721 CTGCAGAGCAGATCGCCTGATGATCCAGCAACCCCCCATCCCGTGGCGCATCTAC 780
Db 3857 CTGCAGAGCAGATCGCCTGATGATCCAGCAACCCCCCATCCCGTGGCGCATCTAC 3916
QY 781 AAGCGGTGATCATCTGGGCTTGAAACAAGATCTGCGATGTACAGCCCGTGAGCATC 840
Db 3917 AAGCGGTGATCATCTGGGCTTGAAACAAGATCTGCGATGTACAGCCCGTGAGCATC 3976
QY 841 CTGCATCAAGCAGAGGCGCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTCAAG 900
Db 3977 CTGCATCAAGCAGAGGCGCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTCTCAAG 4036
QY 901 ACCCTGCGCGCGCGACAGAGCACCAGAGGTGAAGAACTGATGACCGACACCTGCTG 960
Db 4037 ACCCTGCGCGCGCGACAGAGCACCAGAGGTGAAGAACTGATGACCGACACCTGCTG 4096
QY 961 GTGCAGAACGCCAACCCCGACTGCAAGAACCATCTGCGGCTTCGGCCCCGGCGCAGC 1020
Db 4097 GTGCAGAACGCCAACCCCGACTGCAAGAACCATCTGCGGCTTCGGCCCCGGCGCAGC 4156
QY 1021 CTGAGAGAGATGATGACCGCTGCGCAGGCGGTGGCGGCCCAAGAGCGCGCTG 1080
Db 4157 CTGAGAGAGATGATGACCGCTGCGCAGGCGGTGGCGGCCCAAGAGCGCGCTG 4216
QY 1081 CTGCGCGAGGCGAGTGAGCCAGGCGCAACACGAGGTGATGATGAGAAGCAACTTCAAG 1140
Db 4217 CTGCGCGAGGCGAGTGAGCCAGGCGCAACACGAGGTGATGATGAGAAGCAACTTCAAG 4276
QY 1141 GCGCCCCCGCGCATCTGTAAGTGTCTCAACTGCGGCAAGAGGCGCACATCGCCGCAAC 1200
Db 4277 GCGCCCCCGCGCATCTGTAAGTGTCTCAACTGCGGCAAGAGGCGCACATCGCCGCAAC 4336
QY 1201 TGCCCCGCGCCCCCGCAAGAGGCTGTGAAAGTCCGCGCAAGAGGCGCACAGATGAAG 1260
Db 4337 TGCCCCGCGCCCCCGCAAGAGGCTGTGAAAGTCCGCGCAAGAGGCGCACAGATGAAG 4396
QY 1261 GACTGCACCGAGCGCCAGGCCAACTTCTGGGCAAGATCTGGGCCAGCCACAAGGCGCG 1320
Db 4397 GACTGCACCGAGCGCCAGGCCAACTTCTGGGCAAGATCTGGGCCAGCCACAAGGCGCG 4456
QY 1321 CCGGCGCACTTCTCTGCAGAGCGCGCCCGAGGCCACCGCCCCCGCGAGAGCTTCGCG 1380
Db 4457 CCGGCGCACTTCTCTGCAGAGCGCGCCCGAGGCCACCGCCCCCGCGAGAGCTTCGCG 4516
QY 1381 TTGCAAGAGACCAACCCCGGCGCAAGAGCAGAGACCAAGAGCCCTGACCAAGC 1440
Db 4517 TTGCAAGAGACCAACCCCGGCGCAAGAGCAGAGACCAAGAGCCCTGACCAAGC 4576
QY 1441 CTGAAGAGCCTGTTGCGCAACGACCCCTGAGCCATTA 1479
Db 4577 CTGAAGAGCCTGTTGCGCAACGACCCCTGAGCCATTA 4615
```

```
RESULT 10
US-10-190-435-38
; Sequence 38, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYPEPTIDES ENCODING ANTIGENIC HIV TYPE C
; FILE REFERENCE: P18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 4702
; TYPE: DNA
```

ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence:
 OTHER INFORMATION: gp160mod.TV1-gagmod.BW965
 US-10-190-435-38

Query Match 99.9%; Score 1477.4; DB 15; Length 4702;
 Best Local Similarity 99.9%; Pred. No. 9.6e-308;
 Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGGGCGCCCGCGCCAGATCTCGCGCGCGGCAAGCTGACGCGCTGGGAGCGCATCCG 60
Db 3224 ATGGGCGCCCGCGCCAGATCTCGCGCGCGGCAAGCTGACGCGCTGGGAGCGCATCCG 3283
QY 61 CTGCGCGCCCGCGCGCAAGAGTGCTACATGATGAAGCACTGTGTGGGCCAGCCGCGAG 120
Db 3284 CTGCGCGCCCGCGCGCAAGAGTGCTACATGATGAAGCACTGTGTGGGCCAGCCGCGAG 3343
QY 121 CTGAGAAGTTGCGCCCTGAACCCCGGCTGTGAGACCAAGGCGTGCAGAGATC 180
Db 3344 CTGAGAAGTTGCGCCCTGAACCCCGGCTGTGAGACCAAGGCGTGCAGAGATC 3403
QY 181 ATCCGCGAGCTGCACCCCGCCCTGCAGACCGGACGAGAGAGCTGAAGAGCTGTTCAAC 240
Db 3404 ATCCGCGAGCTGCACCCCGCCCTGCAGACCGGACGAGAGAGCTGAAGAGCTGTTCAAC 3463
QY 241 ACCGTGGCCACCCCTGTACTGCGTGCAAGAGATCGAGGTCCGCGCACCAAGAGGCC 300
Db 3464 ACCGTGGCCACCCCTGTACTGCGTGCAAGAGATCGAGGTCCGCGCACCAAGAGGCC 3523
QY 301 CTGACAAGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGATCCAGAGAGGCCGAGGCC 360
Db 3524 CTGACAAGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGATCCAGAGAGGCCGAGGCC 3583
QY 361 GCCGACAAGGGGCAAGGTGAGCCAGACTACCCCATCTGTGAGAACTGTGAGGGCCAGATG 420
Db 3584 GCCGACAAGGGGCAAGGTGAGCCAGACTACCCCATCTGTGAGAACTGTGAGGGCCAGATG 3643
QY 421 GTGCACCAAGGCCATCAGCCCCCGGACCCCTGAACGCTGGGTGAAGGTGATCGAGAGAG 480
Db 3644 GTGCACCAAGGCCATCAGCCCCCGGACCCCTGAACGCTGGGTGAAGGTGATCGAGAGAG 3703
QY 481 GCCTTCAGCCCGGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGGGCCACCCCCAG 540
Db 3704 GCCTTCAGCCCGGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGGGCCACCCCCAG 3763
QY 541 GACCTGAACACGATGTTGAACACCGTGGGGCGGGCCACCAAGCCGCGCATGACATGCTGAAG 600
Db 3764 GACCTGAACACGATGTTGAACACCGTGGGGCGGGCCACCAAGCCGCGCATGACATGCTGAAG 3823
QY 601 GACACCATCAACGAGAGAGGGCCGCGAGTGGGAGCCGCTGCACCCCGTGCACGCGCGGCC 660
Db 3824 GACACCATCAACGAGAGAGGGCCGCGAGTGGGAGCCGCTGCACCCCGTGCACGCGCGGCC 3883
QY 661 ATGCGCCCCCGGCGAGATGCGGAGCGCGCGGCGGAGCGACATCGCGCGGACCAACGAGACC 720
Db 3884 ATGCGCCCCCGGCGAGATGCGGAGCGCGCGGCGGAGCGACATCGCGCGGACCAACGAGACC 3943
QY 721 CTGCAGGAGCAGATCGGCTGTGATGACCAACCCCGCATCCCGTGGGGCGACATCTAC 780
Db 3944 CTGCAGGAGCAGATCGGCTGTGATGACCAACCCCGCATCCCGTGGGGCGACATCTAC 4003
QY 781 AAGCGGTGATCATCTGGGCTGAACAAGATCGTGCGGATGTACAGCCCGGTGAGATC 840
Db 4004 AAGCGGTGATCATCTGGGCTGAACAAGATCGTGCGGATGTACAGCCCGGTGAGATC 4063
QY 841 CTGACATCAAGCAGGGGCCCAAGAGGCCCTTCGCGACTACGTGACCGCTTCTTCAAG 900
Db 4064 CTGACATCAAGCAGGGGCCCAAGAGGCCCTTCGCGACTACGTGACCGCTTCTTCAAG 4123
QY 901 ACCCTGCGCGCGGAGAGACACCCAGAGAGGTGAAGAACTGATGACGACACCCCTGCTG 960
Db 4124 ACCCTGCGCGCGGAGAGACACCCAGAGAGGTGAAGAACTGATGACGACACCCCTGCTG 4183

```

```

QY 961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCGCGCCAGC 1020
Db 4184 GTGCAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCGCGCCAGC 4243
QY 1021 CTGAGAGAGATGATGACCGCTGCGCAGGGCGTGGGGCGGCCAGCCACMAAGCCGCGTG 1080
Db 4244 CTGAGAGAGATGATGACCGCTGCGCAGGGCGTGGGGCGGCCAGCCACMAAGCCGCGTG 4303
QY 1081 CTGCGCGAGGCGATGAGCCAGGCAACACACAGCGTGATGATGACAGAGCAACTTCAAG 1140
Db 4304 CTGCGCGAGGCGATGAGCCAGGCAACACACAGCGTGATGATGACAGAGCAACTTCAAG 4363
QY 1141 GGGCCCCGGCGATGCTCAAGTGTCTCAACTGCGGCAAGGAGGCGCACATCGCCGCAAC 1200
Db 4364 GGGCCCCGGCGATGCTCAAGTGTCTCAACTGCGGCAAGGAGGCGCACATCGCCGCAAC 4423
QY 1201 TGCCGCGCCCCCGCAAGAGGCTGTGAAAGTGCGGCAAGAGGGCCACAGATGAAG 1260
Db 4424 TGCCGCGCCCCCGCAAGAGGCTGTGAAAGTGCGGCAAGAGGGCCACAGATGAAG 4483
QY 1261 GACTGACCCGAGCGCCAGGCCAATTCTTGGGCAAGATCTGCGCCAGCCACAGAGGCGCG 1320
Db 4484 GACTGACCCGAGCGCCAGGCCAATTCTTGGGCAAGATCTGCGCCAGCCACAGAGGCGCG 4543
QY 1321 CCCGGCAACTTCTGACAGAGCGCCCGGAGCCACCGCCCGCGCGGAGAGCTTCCGC 1380
Db 4544 CCCGGCAACTTCTGACAGAGCGCCCGGAGCCACCGCCCGCGCGGAGAGCTTCCGC 4603
QY 1381 TTGAGAGAGACCAACCCCGGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 4604 TTGAGAGAGACCAACCCCGGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4663
QY 1441 CTGAAGAGCTGTTGCGCAACGACCCCTGAGCCAGTAA 1479
Db 4664 CTGAAGAGCTGTTGCGCAACGACCCCTGAGCCAGTAA 4702

```

RESULT 11
 US-10-190-435-17
 Sequence 17, Application US/10190435
 Publication No. US20030143248A1
 GENERAL INFORMATION:
 APPLICANT: ZUR MEGEDE, Jan
 APPLICANT: BARNETT, Susan W.
 APPLICANT: LIAN, Ying
 APPLICANT: ENGELBRECHT, Susan
 APPLICANT: VAN RENSBURG, Estrelita J.
 TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
 FILE REFERENCE: P18133.003 / 2302-18133
 CURRENT APPLICATION NUMBER: US/10/190,435
 CURRENT FILING DATE: 2002-12-30
 NUMBER OF SEQ ID NOS: 319
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 17
 LENGTH: 4716
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence:
 OTHER INFORMATION: GagProteinArtmutatRevNeF_C
 US-10-190-435-17

Query Match 99.9%; Score 1477.4; DB 15; Length 4716;
 Best Local Similarity 99.9%; Pred. No. 9.6e-308;
 Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGGGCGCCCGCGCCAGATCTGCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCGC 60
Db 7 ATGGGCGCCCGCGCCAGATCTGCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCGC 66
QY 61 CTGCGCCCCCGCGCAAGAGTGCTACATGATGAAGCACTGTGTGGGCGAGCCGCGAG 120

```



```
Db      67  CTGCCCCCGCGGCAAGAGTGTCTATGATGAAGCACTGTGTGTGGCCAGCCGCGAG 126
QY      121  CTGAGAAAGTTGCGCTTGAACCCCGGCTGTGTGAGACCAAGGAGGCTGCAAGCAGATC 180
Db      127  CTGAGAAAGTTGCGCTTGAACCCCGGCTGTGTGAGACCAAGGAGGCTGCAAGCAGATC 186
QY      181  ATCCGCGAGCTGCAACCCCGGCTGTGAGACCCGAGCGAGAGCTGAAGAGCTGTTCAC 240
Db      187  ATCCGCGAGCTGCAACCCCGGCTGTGAGACCCGAGCGAGAGCTGAAGAGCTGTTCAC 246
QY      241  ACCGTGGCCACCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACCAAGAGGCC 300
Db      247  ACCGTGGCCACCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACCAAGAGGCC 306
QY      301  CTGAGCAAGATCGAGGAGAGCAGAACTGTCAGCAAGATCCAGCAGAGGCCGAGGCC 360
Db      307  CTGAGCAAGATCGAGGAGAGCAGAACTGTCAGCAAGATCCAGCAGAGGCCGAGGCC 366
QY      361  GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCTGTGCAAGAACTGTGAGGGCCAGATG 420
Db      367  GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCTGTGCAAGAACTGTGAGGGCCAGATG 426
QY      421  GTGCACCAAGGCCATCAGCCCCCGCACCTGAACGCTGGGTGAAGGTGATCGAGAGAA 480
Db      427  GTGCACCAAGGCCATCAGCCCCCGCACCTGAACGCTGGGTGAAGGTGATCGAGAGAA 486
QY      481  GCGTTGAGCCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGGGCCACCCGCCAG 540
Db      487  GCGTTGAGCCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGGGCCACCCGCCAG 546
QY      541  GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAAGGCCCATGCAATGCTGAAG 600
Db      547  GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAAGGCCCATGCAATGCTGAAG 606
QY      601  GACACCATCAACGAGAGAGGCCCGCGAGTGGGACCGCGTGCACCCCGTGCACGCGCGCCC 660
Db      607  GACACCATCAACGAGAGAGGCCCGCGAGTGGGACCGCGTGCACCCCGTGCACGCGCGCCC 666
QY      661  ATGCCCCCGCGGCAAGATGCGCGAGCCCCCGCGGAGCGACATCGCCGCGACCAAGCAAC 720
Db      667  ATGCCCCCGCGGCAAGATGCGCGAGCCCCCGCGGAGCGACATCGCCGCGACCAAGCAAC 726
QY      721  CTGAGAGAGCAGATGCGCTGTGATGACCAACCCGCCATCCCCCGTGGGCGACATCTAC 780
Db      727  CTGAGAGAGCAGATGCGCTGTGATGACCAACCCGCCATCCCCCGTGGGCGACATCTAC 786
QY      781  AAGCGGTGATCATCTGGGCTTGAACAAGATCGTGGGATGTACAGCCCCGTGAGCATC 840
Db      787  AAGCGGTGATCATCTGGGCTTGAACAAGATCGTGGGATGTACAGCCCCGTGAGCATC 846
QY      841  CTGACATCAAGCAGAGGCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAG 900
Db      847  CTGACATCAAGCAGAGGCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAG 906
QY      901  ACCCTGCGCGCGAGCAGACACCAAGAGGTGAAGAACTGATGACCGACACCTGTCTG 960
Db      907  ACCCTGCGCGCGAGCAGACACCAAGAGGTGAAGAACTGATGACCGACACCTGTCTG 966
QY      961  GTGAGAAAGCCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCGCGCGCAGC 1020
Db      967  GTGAGAAAGCCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCGCGCGCAGC 1026
QY      1021  CTGAGAGAGATGATGACCGCTGCAAGGGCGTGGCGGCCCCAGCGCAAGGCCGCGTG 1080
Db      1027  CTGAGAGAGATGATGACCGCTGCAAGGGCGTGGCGGCCCCAGCGCAAGGCCGCGTG 1086
QY      1081  CTGGCGAGGCGATGAGCCAGGCCAACACAGCGTGTATGACAGAAAGCAACTTCAAG 1140
Db      1087  CTGGCGAGGCGATGAGCCAGGCCAACACAGCGTGTATGACAGAAAGCAACTTCAAG 1146
QY      1141  GGCCCCCGCGCATCTCAAGTGTTCAACTGCGGCAAGAGGGCCACATCGCCGCAAC 1200
```

```
Db      1147  GGCCCCCGCGCATCTCAAGTGTTCAACTGCGGCAAGAGGGCCACATCGCCGCAAC 1206
QY      1201  TGCCGCGCCCCCGCAAGAGGCTGTGAAGTCCGCAAGAGGGCCACCAAGATGAAG 1260
Db      1207  TGCCGCGCCCCCGCAAGAGGCTGTGAAGTCCGCAAGAGGGCCACCAAGATGAAG 1266
QY      1261  GACTGACCGAGCGCCAGGCCAACTTCTGGGCAAGATCTGGCCAGCCACAAGGGCCGC 1320
Db      1267  GACTGACCGAGCGCCAGGCCAACTTCTGGGCAAGATCTGGCCAGCCACAAGGGCCGC 1326
QY      1321  CCCGCAACTTCTGTGAGAGCCGCCCCGAGCCACCGCCCCCGCGAGAGCTTCCGC 1380
Db      1327  CCCGCAACTTCTGTGAGAGCCGCCCCGAGCCACCGCCCCCGCGAGAGCTTCCGC 1386
QY      1381  TTCGAGAGACCAACCCCGCGCAAGAGCAGAGCAAGAGACCGCGAGACCTGACAGC 1440
Db      1387  TTCGAGAGACCAACCCCGCGCAAGAGCAGAGCAAGAGACCGCGAGACCTGACAGC 1446
QY      1441  CTGAAGAGCTGTTCGCAACGACCCCTGAGCCAGTAA 1479
Db      1447  CTGAAGAGCTGTTCGCAACGACCCCTGAGCCAGAAA 1485
```

```
RESULT 12
US-10-190-305A-13
; Sequence 13, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; FILE REFERENCE: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/10/190,305A
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 4716
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: GagProteinartmutatRevNef_C
US-10-190-305A-13
```

```
Query Match          99.9%; Score 1477.4; DB 16; Length 4716;
Best Local Similarity 99.9%; Pred. No. 9,6e-308;
Matches 1478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  ATGGGCGCGCGCGCAGCATCTGCGCGCGGCAAGCTGACGCTGTGGAGCGCATCCGC 60
Db      7  ATGGGCGCGCGCGCAGCATCTGCGCGCGGCAAGCTGACGCTGTGGAGCGCATCCGC 66
QY      61  CTGGCCCCCGGCGCAAGAGTGTACATGATGAAGCACTGTGTGGCGCAGCGCGAG 120
Db      67  CTGGCCCCCGGCGCAAGAGTGTGTACATGATGAAGCACTGTGTGGCGCAGCGCGAG 126
QY      121  CTGAGAAAGTTGCGCTTGAACCCCGGCTGTGAGAGCAGGAGGGCTGCAAGCAGATC 180
Db      127  CTGAGAAAGTTGCGCTTGAACCCCGGCTGTGAGAGCAGGAGGGCTGCAAGCAGATC 186
QY      181  ATCCGCGAGCTGACCCCGCTGACAGACCGGCAAGAGAGCTGAAGAGCTGTTCAC 240
Db      187  ATCCGCGAGCTGACCCCGCTGACAGACCGGCAAGAGAGCTGAAGAGCTGTTCAC 246
QY      241  ACCGTGGCCACCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACCAAGAGGCC 300
Db      247  ACCGTGGCCACCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACCAAGAGGCC 306
QY      301  CTGACAAGATCGAGGAGAGCAAGTGTCCAGCAAGATCCAGAGCGGAGGCC 360
```

Db 307 CTGACAAAGATCGAGGAGAGCAGAAACAAGTGCCAGACAAAGATCCAGCAGGCGCCGAGGCC 366
QY 361 GCCGACAAGGGCAAGGTGAGCCAGAACTAATCCCATCGTGCAGAACCTGCAGGGCCAGATG 420
Db 367 GCCGACAAGGGCAAGGTGAGCCAGAACTAATCCCATCGTGCAGAACCTGCAGGGCCAGATG 426
QY 421 GTGACCAAGGCCATCAGCCCCCGCAGCCCTGAACGCTTGGGTGAAGTGTATCGAGAGAAAG 480
Db 427 GTGACCAAGGCCATCAGCCCCCGCAGCCCTGAACGCTTGGGTGAAGTGTATCGAGAGAAAG 486
QY 481 GCCTTCAGCCCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGCGCCACCCCCAG 540
Db 487 GCCTTCAGCCCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGCGCCACCCCCAG 546
QY 541 GACCTGAACACGATGTTGAACAACCGTGGGGCGGCCACCAAGCGCCCATGCAAGATGCTGAAG 600
Db 547 GACCTGAACACGATGTTGAACAACCGTGGGGCGGCCACCAAGCGCCCATGCAAGATGCTGAAG 606
QY 601 GACACCATCAACGAGAGGGCCCGCGAGTGGGACCGCGTGCACCCCGTGCACGCGCGCCCC 660
Db 607 GACACCATCAACGAGAGGGCCCGCGAGTGGGACCGCGTGCACCCCGTGCACGCGCGCCCC 666
QY 661 ATCGCCCCCGGCGCAGATGCGCGAGCGCCCGCGGACGCAATCGCCGCGCACCAACGAGCAC 720
Db 667 ATCGCCCCCGGCGCAGATGCGCGAGCGCCCGCGGACGCAATCGCCGCGCACCAACGAGCAC 726
QY 721 CTGCAAGAGCAGATCGCCTTGATGATGACCAACCCCCCATCCCGTGGGGCGACATCTAC 780
Db 727 CTGCAAGAGCAGATCGCCTTGATGATGACCAACCCCCCATCCCGTGGGGCGACATCTAC 786
QY 781 AAGCGGTGATCATCTGGGGCTGAACAAGATCGTGGCGGATGTACAGCCCCCGTGAGCATC 840
Db 787 AAGCGGTGATCATCTGGGGCTGAACAAGATCGTGGCGGATGTACAGCCCCCGTGAGCATC 846
QY 841 CTGACATCAAGCAGGGCCCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAG 900
Db 847 CTGACATCAAGCAGGGCCCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTTCAAG 906
QY 901 ACCCTGCGCGCGCGAGACAGCAACCCAGAGGTGAAGAACTGGATGACCGACACCTGTCTG 960
Db 907 ACCCTGCGCGCGCGAGACAGCAACCCAGAGGTGAAGAACTGGATGACCGACACCTGTCTG 966
QY 961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCCCGGCGCGAGC 1020
Db 967 GTGCAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCCCGGCGCGAGC 1026
QY 1021 CTGAGAGAGATGATGACCGCCTGCGAGGGCGTGGGGCGGCCCCAGGCCAAGGGCCGCGTG 1080
Db 1027 CTGAGAGAGATGATGACCGCCTGCGAGGGCGTGGGGCGGCCCCAGGCCAAGGGCCGCGTG 1086
QY 1081 CTGGCGGAGGGCGATGAGCCAGGCCAACACCAAGCGTGTATGATGACAGAAAGCAACTTCAAG 1140
Db 1087 CTGGCGGAGGGCGATGAGCCAGGCCAACACCAAGCGTGTATGATGACAGAAAGCAACTTCAAG 1146
QY 1141 GGCCCCCGGCGCATCGTCAAGTGTCTTCAACTGCGGCAAGGAGGGCCCATCGCCCGCAAC 1200
Db 1147 GGCCCCCGGCGCATCGTCAAGTGTCTTCAACTGCGGCAAGGAGGGCCCATCGCCCGCAAC 1206
QY 1201 TGCCGCGCCCCCGCAAGAAAGGCTGTGGAAGTGCGGCAAGGAGGCCACCAAGATGAAG 1260
Db 1207 TGCCGCGCCCCCGCAAGAAAGGCTGTGGAAGTGCGGCAAGGAGGCCACCAAGATGAAG 1266
QY 1261 GACTGCAACGAGCGGCCAACCTTCTGGGCAAGATCTGGCCCAAGGCCCAAGGGCGCG 1320
Db 1267 GACTGCAACGAGCGGCCAACCTTCTGGGCAAGATCTGGCCCAAGGGCGCGCAAGGGCGCG 1326
QY 1321 CCGGCAACTTCTGCAAGAGCGCCCGAGGCCCAACCGCCCCCGCGAGAGCTTCCGC 1380
Db 1327 CCGGCAACTTCTGCAAGAGCGCCCGAGGCCCAACCGCCCCCGCGAGAGCTTCCGC 1386
QY 1381 TTGAGGAGACCAACCCCGCGCAGAAAGAGAGAGCAAGGACCGCGAGACCTTGACAGC 1440
Db 1387 TTGAGGAGACCAACCCCGCGCAGAAAGAGAGAGCAAGGACCGCGAGACCTTGACAGC 1446

QY 1441 CTGAAGAGCCCTGTCGGCAACGACCCCGCTGAGCCAGTAA 1479
Db 1447 CTGAAGAGCCCTGTCGGCAACGACCCCGCTGAGCCAGAAA 1485
RESULT 13
US-10-190-435-20
; Sequence 20, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: P18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 2742
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagTatRevNef_C
US-10-190-435-20

Query Match 99.8%; Score 1475.8; DB 15; Length 2742;
Best Local Similarity 99.9%; Pred. No. 2.2e-307;
Matches 1477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGGCGCGCGCGCAGCATCTGCGCGCGCGCAAGCTGACGCTTGAGCGCATCCGC 60
Db 7 ATGGGCGCGCGCGCAGCATCTGCGCGCGCGCAAGCTGACGCTTGAGCGCATCCGC 66
QY 61 CTGGCGCGCGCGCAAGAGTGTACATGATGAAGCACTGTGTGGCGCGCGAG 120
Db 67 CTGGCGCGCGCGCAAGAGTGTACATGATGAAGCACTGTGTGGCGCGCGAG 126
QY 121 CTGAGAGATTGCGCCTGAACCCCGCGCTGCTGAGACCAAGGAGGCTGCAAGCATC 180
Db 127 CTGAGAGATTGCGCCTGAACCCCGCGCTGCTGAGACCAAGGAGGCTGCAAGCATC 186
QY 181 ATCCGCCAGCTGACCCCGCGCTGACAGACCGGACGAGGAGCTGAAGAGCCTGTTCAAC 240
Db 187 ATCCGCCAGCTGACCCCGCGCTGACAGACCGGACGAGGAGCTGAAGAGCCTGTTCAAC 246
QY 241 ACCGTGGCAACCTGTACTGCGTGACAGAGATCGAGTCCGCGACCAAGAGGCC 300
Db 247 ACCGTGGCAACCTGTACTGCGTGACAGAGATCGAGTCCGCGACCAAGAGGCC 306
QY 301 CTGGAACAAGATGAGAGAGAGAGAAACAAGTCCAGAGATCCAGAGCCGAGGCC 360
Db 307 CTGGAACAAGATGAGAGAGAGAGAAACAAGTCCAGAGATCCAGAGCCGAGGCC 366
QY 361 GCCGACAAGGGGAAGGTGAGCCAGAACTAATCCCATCGTGCAGAACTGCAGGGCCAGATG 420
Db 367 GCCGACAAGGGGAAGGTGAGCCAGAACTAATCCCATCGTGCAGAACTGCAGGGCCAGATG 426
QY 421 GTGCAACGAGGCATCAGCCCCCGCACCTTGAACGCTTGGTGAAGTGTATCGAGAGAAAG 480
Db 427 GTGCAACGAGGCATCAGCCCCCGCACCTTGAACGCTTGGTGAAGTGTATCGAGAGAAAG 486
QY 481 GCCTTCAGCCCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGCGCCACCCCCAG 540
Db 487 GCCTTCAGCCCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGCGCCACCCCCAG 546
QY 541 GACCTGAACACGATGTTGAACAACCGTGGGGCGGCCACCAAGCGCCCATGCAAGATGCTGAAG 600

```
Db 547 GACCTGAACACGATGTTGAACACCGTGGGCGGCCACCAAGGCCCATGTGAGATGCTGAAG 606
QY 601 GACACCATCAACGAGAGAGGCCCGGAGTGGACCGCGGTGCACCCCGGTGCAGCCGGCCCC 660
Db 607 GACACCATCAACGAGAGAGGCCCGGAGTGGACCGCGGTGCACCCCGGTGCAGCCGGCCCC 666
QY 661 ATCGCCCCCGGCGAGATGCGCGAGCCCCCGGCGAGGACATCGCCGCAACCAAGCACC 720
Db 667 ATCGCCCCCGGCGAGATGCGCGAGCCCCCGGCGAGGACATCGCCGCAACCAAGCACC 726
QY 721 CTGCAAGGACAGATGCGCTGTGATGACCAACCCCCCATCCCCGTGGGCGACATCTAC 780
Db 727 CTGCAAGGACAGATGCGCTGTGATGACCAACCCCCCATCCCCGTGGGCGACATCTAC 786
QY 781 AAGCGGTGATCATCTGGGCGCTGAACAAGATCGTGGGATGTACAGCCCCGTGACATC 840
Db 787 AAGCGGTGATCATCTGGGCGCTGAACAAGATCGTGGGATGTACAGCCCCGTGACATC 846
QY 841 CTGACATCAAGCAGGAGCCCCCAAGAGCCCTTCCGCACTACGTGACCGCTTCTTCAAG 900
Db 847 CTGACATCAAGCAGGAGCCCCCAAGAGCCCTTCCGCACTACGTGACCGCTTCTTCAAG 906
QY 901 ACCCTGCGCGCGAGCAGACGACCAAGAGGTGAAGAACTGATGACCGACACCTGTCTG 960
Db 907 ACCCTGCGCGCGAGCAGACGACCAAGAGGTGAAGAACTGATGACCGACACCTGTCTG 966
QY 961 GTGCAAAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTGCGCCCGGCGCAGC 1020
Db 967 GTGCAAAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTGCGCCCGGCGCAGC 1026
QY 1021 CTGAGAGAGATGATGACCGCTGCGCAGGCGCGTGGCGGCCCCAGCCACAAGCGCGCTG 1080
Db 1027 CTGAGAGAGATGATGACCGCTGCGCAGGCGCGTGGCGGCCCCAGCCACAAGCGCGCTG 1086
QY 1081 CTGCGCCGAGGCGGATGAGCCAGGCGCAACCAAGCGGTGATGATGACGAAGAAGCAACTTCAAG 1140
Db 1087 CTGCGCCGAGGCGGATGAGCCAGGCGCAACCAAGCGGTGATGATGACGAAGAAGCAACTTCAAG 1146
QY 1141 GGGCCCCGCGCGCATCGTCAAGTGTCTCAACTGCGGCAAGAGGGGCCACATCGCCGCAAC 1200
Db 1147 GGGCCCCGCGCGCATCGTCAAGTGTCTCAACTGCGGCAAGAGGGGCCACATCGCCGCAAC 1206
QY 1201 TGGCGGCGCCCCCGCAAGAGGGCTGTGGAAGTGGGCAAGAGGGGCCACATCGCCGCAAC 1260
Db 1207 TGGCGGCGCCCCCGCAAGAGGGCTGTGGAAGTGGGCAAGAGGGGCCACATCGCAAG 1266
QY 1261 GACTGCACCCGAGCGCCAGGCCCACTTCTGGGCAAGATCTGCGCCAGCCACAAGGGCCGC 1320
Db 1267 GACTGCACCCGAGCGCCAGGCCCACTTCTGGGCAAGATCTGCGCCAGCCACAAGGGCCGC 1326
QY 1321 CCGGCAACTTCTCTGCAGAGCGCGCCGAGCCCAAGCCCCCGCGCGAGAGCTTCCGC 1380
Db 1327 CCGGCAACTTCTCTGCAGAGCGCGCCGAGCCCAAGCCCCCGCGCGAGAGCTTCCGC 1386
QY 1381 TTGAGAGAGACCAACCCCCCGGCGCAAGACAGAGCAAGAACCCGCGAGACCTTGACCAAGC 1440
Db 1387 TTGAGAGAGACCAACCCCCCGGCGCAAGACAGAGCAAGAACCCGCGAGACCTTGACCAAGC 1446
QY 1441 CTGAAGAGCCTGTTGGCAACGACCCCTGAGCCAGTAA 1479
Db 1447 CTGAAGAGCCTGTTGGCAACGACCCCTGAGCCAGAA 1485
```

RESULT 14
US-10-190-305A-15

```
; Sequence 15, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
```

```
; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190,305A
; CURRENT FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 2742
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagTatRevNef_C
US-10-190-305A-15
```

```
Query Match 99.8%; Score 1475.8; DB 16; Length 2742;
Best Local Similarity 99.9%; Pred. No. 2.2e-307;
Matches 1477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 ATGGGGCGCCGCCAGCATCTGCGGCGGCGCAAGCTGACGCCCTGGAGCGCATCCGC 60
Db 7 ATGGGGCGCCGCCAGCATCTGCGGCGGCGCAAGCTGACGCCCTGGAGCGCATCCGC 66
QY 61 CTGCGCCCCCGGCGCAAGAGTGTCTCATGTGAAGCACTGTGTGGCCAGCCGCGAG 120
Db 67 CTGCGCCCCCGGCGCAAGAGTGTCTCATGTGAAGCACTGTGTGGCCAGCCGCGAG 126
QY 121 CTGGAAGAGTTGCGCTGTAACCCCGGCTGTGAGAGACCAAGGCGTGAAGCAGATC 180
Db 127 CTGGAAGAGTTGCGCTGTAACCCCGGCTGTGAGAGACCAAGGCGTGAAGCAGATC 186
QY 181 ATCCGCAGCTGCACCCCGCTGCAGACCGGCGAGGAGCTGAAGAGCTGTTCAAC 240
Db 187 ATCCGCAGCTGCACCCCGCTGCAGACCGGCGAGGAGCTGAAGAGCTGTTCAAC 246
QY 241 ACCGTGGCACCTGTACTGTGCTGCACGAGAGATGAGGTCCGACACCAAGAGGCC 300
Db 247 ACCGTGGCACCTGTACTGTGCTGCACGAGAGATGAGGTCCGACACCAAGAGGCC 306
QY 301 CTGGAACAAGATGAGAGGAGCAGAACTGCTGAGAGATCCAGCAGCGCGAGGCC 360
Db 307 CTGGAACAAGATGAGAGGAGCAGAACTGCTGAGAGATCCAGCAGCGCGAGGCC 366
QY 361 GCCGACAAAGGCAAGGTGAGCCAGAACTACCCATCTGTGCAAACTGCAAGGCCAGATG 420
Db 367 GCCGACAAAGGCAAGGTGAGCCAGAACTACCCATCTGTGCAAACTGCAAGGCCAGATG 426
QY 421 GTGCACCAAGCCATCAGCCCCCGCACCTGAAAGCCTGGGTGAAGGTGATGAGAGAAG 480
Db 427 GTGCACCAAGCCATCAGCCCCCGCACCTGAAAGCCTGGGTGAAGGTGATGAGAGAAG 486
QY 481 GCCTTCAAGCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCAG 540
Db 487 GCCTTCAAGCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCAG 546
QY 541 GACCTGAACAAGATGTTGAACACCGTGGGCGCCACCAAGCGCCCATGCAATGCTGAAG 600
Db 547 GACCTGAACAAGATGTTGAACACCGTGGGCGCCACCAAGCGCCCATGCAATGCTGAAG 606
QY 601 GACACCATCAACGAGAGAGCGCGCGAGTGGACCGCGTGCACCCCGTGCACCGCGGCC 660
Db 607 GACACCATCAACGAGAGAGCGCGCGAGTGGACCGCGTGCACCCCGTGCACCGCGGCC 666
QY 661 ATCGCCCCCGGCGAGATGCGCGAGCCCCCGGCGAGGACATCGCCGCAACCAAGCACC 720
Db 667 ATCGCCCCCGGCGAGATGCGCGAGCCCCCGGCGAGGACATCGCCGCAACCAAGCACC 726
QY 721 CTGCAAGGACAGATGCGCTGTGATGACCAACCCCCCATCCCCGTGGGCGACATCTAC 780
Db 727 CTGCAAGGACAGATGCGCTGTGATGACCAACCCCCCATCCCCGTGGGCGACATCTAC 786
QY 781 AAGCGGTGATCATCTGGGCGCTGAACAAGATCGTGGGATGTACAGCCCCGTGAGCATC 840
Db 787 AAGCGGTGATCATCTGGGCGCTGAACAAGATCGTGGGATGTACAGCCCCGTGAGCATC 846
```

```
QY      841 CTGGACATCAAGGAGGGCCCCCAAGAGAGCCCTTCCCGGACTACGTGGAACCGCTTCTTCAAG 900
      |||
Db      847 CTGGACATCAAGGAGGGCCCCCAAGAGAGCCCTTCCCGGACTACGTGGAACCGCTTCTTCAAG 906
QY      901 ACCCTGGCGCGCGAGCAAGAGCAACCCAGAGAGGTGAAGAAGTGAATGATGACCGACACCTGCTG 960
      |||
Db      907 ACCCTGGCGCGCGAGCAAGAGCAACCCAGAGAGGTGAAGAAGTGAATGATGACCGACACCTGCTG 966
QY      961 GTGAGAGAACGCCAACCCCGGACTGCAAGAACCATCTCTCGCGGCTCTTGGGCCCCGGCGCCAGC 1020
      |||
Db      967 GTGAGAGAACGCCAACCCCGGACTGCAAGAACCATCTCTCGCGGCTCTTGGGCCCCGGCGCCAGC 1026
QY      1021 CTGAGAGAGATGATGACCGGCTGCGAGGGCGGTGGGGGGCCCCCAGCCACAAGGCCCGCTG 1080
      |||
Db      1027 CTGAGAGAGATGATGACCGGCTGCGAGGGCGGTGGGGGGCCCCCAGCCACAAGGCCCGCTG 1086
QY      1081 CTGGCCGAGGCGGATGAGCCAGGCCACAACCAAGCGTGATGATGCAGAAAGAGCAACTTCAAG 1140
      |||
Db      1087 CTGGCCGAGGCGGATGAGCCAGGCCACAACCAAGCGTGATGATGCAGAAAGAGCAACTTCAAG 1146
QY      1141 GGGCCCCCGGCGCATCTCAAGTGTCTTCAACTGCGGCAAGAGGGCCACATCGCCCGCAAC 1200
      |||
Db      1147 GGGCCCCCGGCGCATCTCAAGTGTCTTCAACTGCGGCAAGAGGGCCACATCGCCCGCAAC 1206
QY      1201 TGCCGCGCCCCCGGCAAGAGGGGCTGCTGGAAGTGGCGGCAAGAGGGCCACCATGAAG 1260
      |||
Db      1207 TGCCGCGCCCCCGGCAAGAGGGGCTGCTGGAAGTGGCGGCAAGAGGGCCACCATGAAG 1266
QY      1261 GACTGCACCCGAGCGGCCAGGCCCAACTTCTTGGGCAAGATCTGGCCCAACCAAGGGCGCG 1320
      |||
Db      1267 GACTGCACCCGAGCGGCCAGGCCCAACTTCTTGGGCAAGATCTGGCCCAACCAAGGGCGCG 1326
QY      1321 CCGGCACTTCTCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
      |||
Db      1327 CCGGCACTTCTCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1386
QY      1381 TTGAGGAGAGCAACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
      |||
Db      1387 TTGAGGAGAGCAACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1446
QY      1441 CTGAAGAGCCTGTTGGCAACGAGACCCCTTGAGCCAGTAA 1479
      |||
Db      1447 CTGAAGAGCCTGTTGGCAACGAGACCCCTTGAGCCAGAA 1485
```

```
RESULT 15
US-10-190-435-9
; Sequence 9, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEDEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 3930
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagComplPolmut_C
US-10-190-435-9
```

```
Query Match          99.8%; Score 1475.8; DB 15; Length 3930;
Best Local Similarity 99.9%; Pred. No. 2.1e-307;
Matches 1477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 ATGGCGCGCGCGCCAGCATCTGCGCGGCGGCAAGCTGGAACGCTTGGAGCGCATCCGC 60
      |||
Db      7 ATGGCGCGCGCGCCAGCATCTGCGCGGCGGCAAGCTGGAACGCTTGGAGCGCATCCGC 66
QY      61 CTGCGCGCGCGCGCAAGAGTGTCTACATGATGAAGCACTTGTGTGGGCGAGCCGCGAG 120
      |||
Db      67 CTGCGCGCGCGCGCAAGAGTGTCTACATGATGAAGCACTTGTGTGGGCGAGCCGCGAG 126
QY      121 CTGAGAGAGTTGCGCTTGAACCCCGGCTGTGAGAGACCAAGGAGGCTGCAAGCAGATC 180
      |||
Db      127 CTGAGAGAGTTGCGCTTGAACCCCGGCTGTGAGAGACCAAGGAGGCTGCAAGCAGATC 186
QY      181 ATCCGCCAGCTGCACCCCGGCTGCAGACCGGCGAGGAGAGCTGAAGAGCTGTTCAAC 240
      |||
Db      187 ATCCGCCAGCTGCACCCCGGCTGCAGACCGGCGAGGAGAGCTGAAGAGCTGTTCAAC 246
QY      241 ACCGTGGCCACCTGTACTGCGTGCAAGAGAGATGAGGTCCGCGACACCAAGAGGCGC 300
      |||
Db      247 ACCGTGGCCACCTGTACTGCGTGCAAGAGAGATGAGGTCCGCGACACCAAGAGGCGC 306
QY      301 CTGAGCAAGATCGAGAGAGAGAGCAAAAGTGCCAGCAGAAATCCAGAGGCGGAGGCC 360
      |||
Db      307 CTGAGCAAGATCGAGAGAGAGAGCAAAAGTGCCAGCAGAAATCCAGAGGCGGAGGCC 366
QY      361 GCCGACAAAGGCAAGGTGAGCCAGAACTACCCCATCTGTGCAAACTGCAAGGCGCAGATG 420
      |||
Db      367 GCCGACAAAGGCAAGGTGAGCCAGAACTACCCCATCTGTGCAAACTGCAAGGCGCAGATG 426
QY      421 GTGCACCAAGGCGATCAGCCCCCGCACCTGTAACGCTGGGTGAAGGTGATCGAGAGAA 480
      |||
Db      427 GTGCACCAAGGCGATCAGCCCCCGCACCTGTAACGCTGGGTGAAGGTGATCGAGAGAA 486
QY      481 GCCTTCAGCCCGGAGGTGATCCCATGTTCAACGCGCTGAGCGAGGCGGCAACCCCGCAG 540
      |||
Db      487 GCCTTCAGCCCGGAGGTGATCCCATGTTCAACGCGCTGAGCGAGGCGGCAACCCCGCAG 546
QY      541 GACCTGAACAGATGTTGAACAACGTGGGCGGCAACCAAGCGCCCATGCAATGCTGAAG 600
      |||
Db      547 GACCTGAACAGATGTTGAACAACGTGGGCGGCAACCAAGCGCCCATGCAATGCTGAAG 606
QY      601 GACACCATCAACGAGAGGCGCGCGAGTGGACCGGTCACCCCGTGACGCGCGCGCGCG 660
      |||
Db      607 GACACCATCAACGAGAGGCGCGCGAGTGGACCGGTCACCCCGTGACGCGCGCGCGCGCG 666
QY      661 ATCGCCCCCGGCGAGATGCGGCGAGCCCCCGGCGAGCGACATCGCGCGCACCAAGCAGCACC 720
      |||
Db      667 ATCGCCCCCGGCGAGATGCGGCGAGCCCCCGGCGAGCGACATCGCGCGCACCAAGCAGCACC 726
QY      721 CTGAGAGAGATCGCCTGATGACCAAGACCCCGCATCCCGTGGGCGACATCTAC 780
      |||
Db      727 CTGAGAGAGATCGCCTGATGACCAAGACCCCGCATCCCGTGGGCGACATCTAC 786
QY      781 AAGCGGTGATCATCTGGGCTGAACAAGATCGTGCGATGTACAGCCCCGTGAGCATC 840
      |||
Db      787 AAGCGGTGATCATCTGGGCTGAACAAGATCGTGCGATGTACAGCCCCGTGAGCATC 846
QY      841 CTGACATCAAGCAGAGGCGCGCAAGAGGCGCTTCCGCGACTACGTGAGCCGCTTCTCAAG 900
      |||
Db      847 CTGACATCAAGCAGAGGCGCGCAAGAGGCGCTTCCGCGACTACGTGAGCCGCTTCTCAAG 906
QY      901 ACCCTGGCGCGCGAGCAGAGACCCAGAGGTGAAGAAGTGAATGACCGACACCTGCTG 960
      |||
Db      907 ACCCTGGCGCGCGAGCAGAGACCCAGAGGTGAAGAAGTGAATGACCGACACCTGCTG 966
QY      961 GTGCAAGAACGCCAACCCCGACTGCAAGACCATCTTGCGGCTCTCGGCCCCGGCGCCAGC 1020
      |||
Db      967 GTGCAAGAACGCCAACCCCGACTGCAAGACCATCTTGCGGCTCTCGGCCCCGGCGCCAGC 1026
QY      1021 CTGAGAGAGATGATGACCGGCTGCGAGGGCGGTGGGGCGGCCCCAGCACCAAGGCCCGGCTG 1080
      |||
Db      1027 CTGAGAGAGATGATGACCGGCTGCGAGGGCGGTGGGGCGGCCCCAGCACCAAGGCCCGGCTG 1086
```

QY 1081 CTGCGCGAGGCGATGAGCCAGGCCAACACCAAGCGTGATGATGCAGAAAGACCAACTTCAAG 1140
|||
Db 1087 CTGCGCGAGGCGATGAGCCAGGCCAACACCAAGCGTGATGATGCAGAAAGACCAACTTCAAG 1146
QY 1141 GGCCCCCGGCGCATCGTCAAGTGCTTCAACTGCGGCAAGAGGGCCACATGCCCCGCAAC 1200
|||
Db 1147 GGCCCCCGGCGCATCGTCAAGTGCTTCAACTGCGGCAAGAGGGCCACATGCCCCGCAAC 1206
QY 1201 TGCCGCGCCCCCGCGCAAGAGGGCTGTGAAAGTGCGGCAAGAGGGCCACAGATGAAG 1260
|||
Db 1207 TGCCGCGCCCCCGCGCAAGAGGGCTGTGAAAGTGCGGCAAGAGGGCCACAGATGAAG 1266
QY 1261 GACTGCAACCGAGCGCGCCAGGCCCAACTTCTGGGCAAGATCTGGCCCAAGGCCCAAGGGCCGC 1320
|||
Db 1267 GACTGCAACCGAGCGCGCCAGGCCCAACTTCTGGGCAAGATCTGGCCCAAGGCCCAAGGGCCGC 1326
QY 1321 CCGGCAACTTCTGCAAGAGCGCGCCGAGCCCAACCGCCCCCGCGAGAGCTTCCGC 1380
|||
Db 1327 CCGGCAACTTCTGCAAGAGCGCGCCGAGCCCAACCGCCCCCGCGAGAGCTTCCGC 1386
QY 1381 TTCGAGAGACCAACCCCGGCGCAGAAAGAGAGCAAGACCGCGACACCTGACCAAC 1440
|||
Db 1387 TTCGAGAGACCAACCCCGGCGCAGAAAGAGAGCAAGACCGCGACACCTGACCAAC 1446
QY 1441 CTGAAGAGCTGTTCGGCAACGACCCCTGAGCCAGTAA 1479
|||
Db 1447 CTGAAGAGCTGTTCGGCAACGACCCCTGAGCCAGTAA 1485

Search completed: September 1, 2005, 23:17:17
Job time : 920.198 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 1, 2005, 21:17:23 ; Search time 25.739 Seconds
(without alignments)
8578.900 Million cell updates/sec

Title: US-09-475-704A-3
Perfect score: 2772
Sequence: 1 atggggcgccgcgcgcagcat.....acgaccccttgagccagtaa 1479

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 74649064 residues
Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09475704/runat_01092005_160751_19541/app_query.fasta_1.3342
-DB=Issued_Patents_AA_QFMT=faetan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09475704 @CGN_1_1_53 @runat_01092005_160751_19541 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCFUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2435	87.8	492	4	US-09-991-258-5 Sequence 5, Appli
2	2256	81.4	500	2	US-08-816-155B-45 Sequence 45, Appli
3	2256	81.4	500	3	US-09-079-587-45 Sequence 45, Appli
4	2249	81.1	500	2	US-08-375-510-1 Sequence 1, Appli
5	2249	81.1	500	2	US-08-487-657-1 Sequence 1, Appli
6	2249	81.1	500	4	US-09-309-572-16 Sequence 16, Appli
7	2249	81.1	500	4	US-09-718-096-16 Sequence 16, Appli
8	2245	81.0	1350	4	US-09-952-060-35 Sequence 35, Appli
9	2235	80.6	512	3	US-08-463-210-8 Sequence 8, Appli
10	2235	80.6	512	3	US-09-124-900-2 Sequence 2, Appli
11	2235	80.6	512	4	US-08-463-028-8 Sequence 8, Appli
12	2106	76.0	493	4	US-09-952-060-33 Sequence 33, Appli

13	2037.5	73.5	478	1	US-08-127-499A-11	Sequence 11, Appli
14	2037.5	73.5	478	1	US-08-482-847-11	Sequence 11, Appli
15	2023	73.0	437	3	US-08-392-794A-2	Sequence 2, Appli
16	1976	71.3	512	4	US-09-319-588C-4	Sequence 4, Appli
17	1797.5	64.8	498	1	US-08-470-202-59	Sequence 59, Appli
18	1797.5	64.8	498	1	US-08-471-770-59	Sequence 59, Appli
19	1797.5	64.8	498	2	US-08-468-059-59	Sequence 59, Appli
20	1797.5	64.8	498	3	US-09-109-916-59	Sequence 59, Appli
21	1797.5	64.8	498	4	US-09-886-156-59	Sequence 59, Appli
22	1797.5	64.8	498	4	US-09-886-149-59	Sequence 59, Appli
23	1797.5	64.8	498	4	US-09-886-150-59	Sequence 59, Appli
24	1797.5	64.8	498	4	US-09-886-159-59	Sequence 59, Appli
25	1797.5	64.8	498	4	US-10-326-090-59	Sequence 59, Appli
26	1775.5	64.1	498	1	US-08-470-202-60	Sequence 60, Appli
27	1775.5	64.1	498	1	US-08-471-770-60	Sequence 60, Appli
28	1775.5	64.1	498	2	US-08-468-059-60	Sequence 60, Appli
29	1775.5	64.1	498	3	US-09-109-916-60	Sequence 60, Appli
30	1775.5	64.1	498	4	US-09-886-156-60	Sequence 60, Appli
31	1775.5	64.1	498	4	US-09-886-149-60	Sequence 60, Appli
32	1775.5	64.1	498	4	US-09-886-150-60	Sequence 60, Appli
33	1775.5	64.1	498	4	US-09-886-159-60	Sequence 60, Appli
34	1775.5	64.1	498	4	US-10-326-090-60	Sequence 60, Appli
35	1681	60.6	363	2	US-08-850-049-130	Sequence 130, App
36	1681	60.6	363	2	US-08-050-478-130	Sequence 130, App
37	1681	60.6	363	3	US-09-414-117-130	Sequence 130, App
38	1681	60.6	363	4	US-09-678-437-130	Sequence 130, App
39	1681	60.6	363	4	US-09-943-722-130	Sequence 130, App
40	1536.5	55.4	458	1	US-07-648-796A-5	Sequence 5, Appli
41	1536.5	55.4	799	1	US-07-648-796A-7	Sequence 7, Appli
42	1530.5	55.2	337	1	US-07-648-796A-1	Sequence 1, Appli
43	1514.5	54.6	328	4	US-08-776-188C-77	Sequence 77, Appli
44	1478.5	53.3	518	4	US-09-206-551-45	Sequence 45, Appli
45	1413	51.0	294	3	US-09-370-368-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-991-258-5
Sequence 5, Application US/09991258
Patent No. 6783939
GENERAL INFORMATION:
APPLICANT: Olmsted, Robert
APPLICANT: Keith, Paula
APPLICANT: Dryga, Sergey
APPLICANT: Caley, Ian
APPLICANT: Maughan, Maureen
APPLICANT: Johnston, Robert
APPLICANT: Davis, Nancy
APPLICANT: Swastrom, Ronald
TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOMES WITH MODIFIED HIV GENES FOR USE
FILE REFERENCE: 0113.0001U3
CURRENT APPLICATION NUMBER: US/09/991,258
CURRENT FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: 09/902,537
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: 60/216,995
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 492
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence; No. 6783939e =
US-09-991-258-5
Alignment Scores:
Pred. No.: 2.97e-141 Length: 492
Score: 2435.00 Matches: 458

Percent Similarity:	96.96%	Conservative:	20
Best Local Similarity:	92.90%	Mismatches:	13
Query Match:	87.84%	Indels:	2
DB:	4	Gaps:	2

QY	1	ATGGGCGCCGCGCGCAGCATCCTGTGGCGGGCGGCAAGCTGGACGCTGGGAGCGCATCCGC	60
Db	1	MetAlaAlaArgAlaSerIleLeuArgGlyGlyLeuValLeuAspLysTrpGlyLysIleArg	20
QY	61	CTGGCGCCCGCGCGCAGAGTGTACATGATGATGAAGCACTGTGTGGGCCAGCGCGGAG	120
Db	21	LeuArgProGlyGlyLysLysHisTyrMetLeuYHisIleValTrpAlaSerArgGly	40
QY	121	CTGGAGAAGTTCGCTGTAAACCCCGGCTGTGGAGACCCAGCGGCTGTCAAGCATC	180
Db	41	LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrSerGluGlyCysLysGlnIle	60
QY	181	ATCCGCGCAGCTGCACCCCGCCTGCAGACCCGCGCAGGAGCTGAAGAGCCTTTCAAC	240
Db	61	MetLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluLeuLysSerLeuTyrAsn	80
QY	241	ACCGTGGCCACCTGTACTGCGTGACAGAAAGATCGAGTCCGCGCACCAAGAGGCC	300
Db	81	ThrValAlaThrLeuTyrCysValHisGlyLysIleGluValArgAspThrLysGluAla	100
QY	301	CTGCACAAGATCGAGAGGAGGACAGAACAGTCCAGCAAGAGATCCAGCGCGCAGGCC	360
Db	101	LeuAspLysIleGluGluGluGlnAsnLysCysGlnGlnLysThrGlnGlnAlaLysAla	120
QY	361	GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATGTCAGAACTGCAGGGCCAGATG	420
Db	121	AlaAsp--GlyLysValSerGlnAsnTyrProIleValGlnAsnLeuGlnGlyMet	139
QY	421	GTCGACCAGGCCATCAGCCCCCGCCTGAACGCTGGTGAAGTGTATCGAGAGAAG	480
Db	140	ValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValIleGluGlnLys	159
QY	481	GCTTCAGCCCGGAGGTGATCCCATGTTCAACGCGCTGAGCGAGGGCGCACCCCCAG	540
Db	160	AlaPheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGln	179
QY	541	GACCTGAACACGATGTTGAACACCGGTGGCGGCCACAGCGCCATGCAGATGCTGAAG	600
Db	180	AspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLys	199
QY	601	GACACCATCAACGAGAGGCGCGCGAGTGGAGCCGCGTGCACCCCGTGCAGCGCGGCC	660
Db	200	AspThrIleAsnGluGlnAlaIaIaGluTrpAspArgLysIleHisProValHisAlaGlyPro	219
QY	661	ATCGCCCCCGCGCAGATGCGCGAGCCCCGCGCGCAGCATCGCCGGCACCAACGACACC	720
Db	220	IleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerThr	239
QY	721	CTGCAGAGCAGATCGCCTGATGACAGCAACCCCCCATCCCGCTGGCGCAGATCTAC	780
Db	240	LeuGlnGluGlnIleAlaTrpMetThrSerAsnProIleProValGlyAspIleTyr	259
QY	781	AAGCGGTGATCATCCTGGGCTGAACAAGATCGTGGGATGTACAGCCCGTGAGCATC	840
Db	260	LysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProValSerIle	279
QY	841	CTGCAGATCAAGCAGGGGCCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAG	900
Db	280	LeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPheHelys	299
QY	901	ACCCTGCGCGCGCAGACAGACCCAGAGAGGTGAAGAACTGATGACCGACACCTGCTG	960
Db	300	ThrLeuArgAlaGluGlnAlaThrGlnGluValLysAsnTrpMetThrAspThrLeuLeu	319
QY	961	GTCGAGAAGCGCAACCCCGACTGCAAGACCATCCTGGCGGCTTCGGCCCCGGCGGCAGC	1020

```

Db      320 ValGlnAsnAlaAsnProAspCysLysThrIleLeuAlaLeuGlyProGlyAlaThr 339
Qy      1021 CTGAGAGATGATGACCGCCCTGCGAGGGCGTGCGGCGCCAGCCACAAGGCCCGCGTG 1080
Db      340 LeuGlnGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLysAlaArgVal 359
Qy      1081 CTGCGCCGAGCGCATGAGCCAGGCCCAACACC--AGCGTATGATGCAGAAAGCAACTTC 1137
Db      360 LeuAlaGlnAlaMetSerGlnThrAsnSerGlyAsnIleMetMetGlnArgSerAsnIle 379
Qy      1138 AAGGGCCCCCGCGCATCGTCAAGTGTCTTCAACTGCGGCAAGAGGGCCACATGCGCCGC 1197
Db      380 LysGlyProAlaArgAlaValLysCysPheAsnCysGlyLysGlnGlyHisIleAlaArg 399
Qy      1198 AACTGCGCGCGCCCCCGCAGAAAGGGCTGCTGGAAGTGCGGCAAGAGGGCCACCAGATG 1257
Db      400 AsnCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGlnGlyHisGlnMet 419
Qy      1258 AAGGACTGCACCGAGCGCGCCAGGCCCAACTTCTGTGGGCAAGATCTGGGCCAGCCACAAGGGC 1317
Db      420 LysAspCysThrGlnArgGlnAlaAsnIleGlnGlyLysIleTrpProSerHisLysGly 439
Qy      1318 CGCCCCCGCAACTTCTCTGAGAGCGCGCGCCGAGAGCCAGCCAGCCCGCGGAGAGCTTC 1377
Db      440 ArgProGlyAsnIlePheLeuGlnAsnArgProGluProThrAlaProAlaGluSerPhe 459
Qy      1378 CGCTTCGAGGAGACCAACCCCGCGCCAGAGCAGAGCAGAGCAAGACCGCGAGACCTTGAC 1437
Db      460 ArgPheGlnGluThrThrProAlaProLysGlnGluProIleGlnArgGluProLeuThr 479
Qy      1438 AGCCTGAAGACCTGTTCGCGCAACGACCCCTGAGCCAG 1476
Db      480 SerLeuLysSerLeuPheGlySerAspProLeuSerGln 492

RESULT 2
US-08-816-155B-45
; Sequence 45, Application US/08816155B
; Patent No. 5990091
; GENERAL INFORMATION:
; APPLICANT: TARTAGLIA, JAMES
; APPLICANT: COX, WILLIAM I.
; APPLICANT: GETTIG, RUSSELL R.
; APPLICANT: MARTINEZ, HECTOR
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: PINCUS, STEVEN E.
; TITLE OF INVENTION: VECTORS HAVING ENHANCED EXPRESSION, AND
; TITLE OF INVENTION: METHODS OF MAKING AND USES THEREOF
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP
; STREET: 745 FIFTH AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10151
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/816,155B
; FILING DATE: 12-MAR-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: KOWALSKI, THOMAS J.
; REGISTRATION NUMBER: 32,147
; REFERENCE/DOCKET NUMBER: 454310-2990
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-588-0800
; TELEFAX: 212-588-0500
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:

```



```

; LENGTH: 500 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-816-155B-45

```

Alignment Scores:

Pred. No.:	2.59e-130	Length:	500
Score:	2256.00	Matches:	424
Percent Similarity:	92.42%	Conservative:	39
Best Local Similarity:	84.63%	Mismatches:	28
Query Match:	81.39%	Indels:	10
DB:	2	Gaps:	5

US-09-475-704A-3 (1-1479) x US-08-816-155B-45 (1-500)

```

QY 1 ATGGGCGCCCGCCAGCATCTCGCGCGCGCAAGCTGAGCGCCTGGAGAGCGCATCCGC 60
   |||||
Db 1 MetGlyAlaArgAlaSerValLeuSerGlyGlyLeuAspArgTyrGlyLysIleArg 20

QY 61 CTGCGCCCGCGCGCGCAAGAGTGTACATGATGAAGCACTGTGTGGCCAGCCCGCAG 120
   |||||
Db 21 LeuArgProGlyGlyLysLysLysTyrLysLeuLysHisIleValTyrAlaSerArgGlu 40

QY 121 CTGAGAGAAGTTGCGCCTGAACCCCGCGCTGTGAGAGCAGCGAGGCTGCAAGCAGATC 180
   |||||
Db 41 LeuGluArgPheAlaValAsnProGlyLeuLeuGluThrSerGluGlyCysArgGlnIle 60

QY 181 ATCCGCGCAGCTGCACCCCGCGCTGCAGACCGCGAGGAGAGCTGAAGACCTGTTCAAC 240
   |||||
Db 61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGluGlnLeuArgSerLeuTyrAsn 80

QY 241 ACCGTGGCCACCCCTGTACTGCGTGCAAGAGATCGAGGTCCGCGACCAAGAGAGGCC 300
   |||||
Db 81 ThrValAlaThrLeuTyrCysValHisGlnArgIleGluIleLysAspThrLysGlnAla 100

QY 301 CTGAGACAAGATCGAGAGGAGCAGAAAGTGCACGAGAGATCCAGAGCGCGAGGCC 360
   |||||
Db 101 LeuAspLysIleGluGluGlnGlnAsnLysSerLysLysAlaGlnGlnAla--Ala 119

QY 361 GCCGACAAAGGC-----AAGTGAGCCAGAACTACCCCATCGTGCAAGCCTGCAG 411
   |||||
Db 120 AlaAspThrGlyHisSerAsnGlnValSerGlnAsnTyrProIleValGlnAsnIleGln 139

QY 412 GGCCAGATGTGCACAGGCCCATCAGCCCGCGACCTGAAGCGCTGGTGAAGTGATC 471
   |||||
Db 140 GlyGlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTyrValLysValVal 159

QY 472 GAGGAGAAGGCTTCAGACCCCGAGGTATCCCATGTTCAACCGCCTGAGCGAGGCC 531
   |||||
Db 160 GluGluLysAlaPheSerProGluValIleProMetPheSerAlaLeuSerGluGlyAla 179

QY 532 ACCCCCGAGGACCTGAACACGATGTTGAACACCGTGCGCGCCAGCAGCCCGCATGCAG 591
   |||||
Db 180 ThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlnHisGlnAlaAlaMetGln 199

QY 592 ATGCTGAAGGACACCATCAACGAGAGGCGCGCGAGTGAGCCGCGTGACCCCGTGAC 651
   |||||
Db 200 MetLeuLysGluThrIleAsnGluGlnAlaAlaGluTyrPaspArgValHisProValHis 219

QY 652 GCCGCGCCCATCGCCCGCGCGCAGATGCGAGCGCGCGCGCAGCGCATCGCCGCGACC 711
   |||||
Db 220 AlaGlyProIleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThr 239

QY 712 ACCAGGACCTGCAGAGCAGATGCGCTGATGACCAACCCCGCATCCCGGTGGGC 771
   |||||
Db 240 ThrSerThrLeuGlnGluGlnIleGlyTyrMetThrAsnAsnProIleProValGly 259

QY 772 GACATCTACAAGCGGTGATCATCTCTGGGCTGAACAAGATCGTGCGGATGTACAGCCC 831
   |||||
Db 260 GluIleTyrLysArgTyrIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerPro 279

```

```

QY 832 GTGAGCATCTTGACATCAAGCAGGCGCCCAAGAGACCTTCCGCACTACGTGACCGC 891
   |||||
Db 280 ThrSerIleLeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArg 299

QY 892 TTCCTCAAGACCTTCGCGCGCGCAGAGCAGCACCAGAGGTGAAGAACTGATGACCGAC 951
   |||||
Db 300 PheTyrLysThrLeuArgAlaGluGlnAlaSerGlnGluValLysAsnTyrMetThrGlu 319

QY 952 ACCCTGTGTGTCAGAACGCCAACCCGCACTGCAAGACCATCTGCGCTCGGCCCC 1011
   |||||
Db 320 ThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyPro 339

QY 1012 GCGCGCAGCCTGAGAGATGATGACCGCCTGCGCAGGCGGTGGCGGCCAGCAAG 1071
   |||||
Db 340 AlaAlaThrLeuGluGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLys 359

QY 1072 GCCCGCTGTGCGCGCAGGCGATGAGCCAGGCCAACACAGC-----GTGATGTCAG 1125
   |||||
Db 360 AlaArgValLeuAlaGluAlaMetSerGlnValThrAsnSerAlaThrIleMetGln 379

QY 1126 AAGACCACTTCAAGGCGCCCGCGCATGTCAGTGTCTCAACTGCGCGCAAGAGGCC 1185
   |||||
Db 380 ArgGlyAsnPheArgAsnGlnArgLysIleValLysCysPheAsnCysGlyLysGlu 399

QY 1186 CACATCGCCCGCAACTGCGCGCGCGCGCGCGCAAGAGGCTGTGGAAGTGGCGCAAGAG 1245
   |||||
Db 400 HisThrAlaArgAsnCysArgAlaProArgLysLysGlyCysTyrLysCysGlyLysGlu 419

QY 1246 GGCCACCAAGATGAAGACTGCACCGAGCGCCAGGCCAATTCTTGGGCAAGATCTGGCCC 1305
   |||||
Db 420 GlyHisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTyrPro 439

QY 1306 AGCCACAAGGCGCGCGCGCACTTCTGCAAGACCGCGCGCGCAGCCACCGCCCCCCC 1365
   |||||
Db 440 SerTyrLysGlyArgProGlyAsnPheLeuGlnSerArgProGluProThrAlaProPro 459

QY 1366 GCCGAGAGCTTCCGC-----TTCGAGGAGACCAACCCCGCGCGCAAGAGAGCAAG 1419
   |||||
Db 460 GluGluSerPheArgSerGlyValGluThrThrProProGlnLysGlnGluProIle 479

QY 1420 GACCGCGAG-----ACCTGACCAAGCCTGAAGAGCCTGTTCCGCAACGACCCCTGAGC 1473
   |||||
Db 480 AspLysGluLeuTyrProLeuThrSerLeuArgSerLeuPheGlyAsnAspProSerSer 499

QY 1474 CAG 1476
   |||
Db 500 Gln 500

RESULT 3
US-09-079-587-45
; Sequence 45, Application US/09079587
; Patent No. 6130066
; GENERAL INFORMATION:
; APPLICANT: TARTAGLIA, JAMES
; APPLICANT: COX, WILLIAM I.
; APPLICANT: GETTIG, RUSSELL R.
; APPLICANT: MARTINEZ, HECTOR
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: PINCUS, STEVEN E.
; TITLE OF INVENTION: VECTORS HAVING ENHANCED EXPRESSION, AND
; TITLE OF INVENTION: METHODS OF MAKING AND USES THEREOF
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP
; STREET: 745 FIFTH AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10151
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/079,587
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/816,155
; FILING DATE: 12-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: KOWALSKI, THOMAS J.
; REGISTRATION NUMBER: 32,147
; REFERENCE/DOCKET NUMBER: 454310-2990
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-588-0800
; TELEFAX: 212-588-0500
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 500 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-09-079-587-45

Alignment Scores:
Pred. No.: 2.59e-130 Length: 500
Score: 2256.00 Matches: 424
Percent Similarity: 92.42% Conservative: 39
Best Local Similarity: 84.63% Mismatches: 28
Query Match: 81.39% Indels: 10
DB: 3 Gaps: 5

US-09-475-704A-3 (1-1479) x US-09-079-587-45 (1-500)

QY 1 ATGGGCGCCCGCGCCAGCATCTCGCGGCGGCAAGCTGGACGCGCTGGGAGCGCATCCGC 60
Db 1 MetGlyAlaArgAlaSerValLeuSerGlyGlyGluLeuAspArgTirpGluLysIleArg 20
QY 61 CTGGCGCCCGCGGCGCAAGAAGTGTACATGATGAAGCACTGTGTGGGCCAGCCGCGAG 120
Db 21 LeuArgProGlyGlyLysLysLysTyrLysLeuLysHisIleValTirpAlaSerArgGlu 40
QY 121 CTGGAGAAGTTCGCGCTGAACCCCGGCTGTGAGACCAGCGAGGCGTGCAGACGATC 180
Db 41 LeuGluArgPheAlaValAsnProGlyLeuLeuGluThrSerGluGlyCysArgGlnIle 60
QY 181 ATCCGCGCAGCTGCACCCCGCCCTGCAGACCGGACGAGGAGGCTGAAGACCTGTTCAC 240
Db 61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGluGluLeuArgSerLeuTyrAsn 80
QY 241 ACCGTGGCCACCCCTGTACTGCGTGCAGAGAAAGATCGAGGTCGCGCAACCAAGAGGCC 300
Db 81 ThrValAlaThrLeuTyrCysValHisGlnArgIleGluIleLysAspThrLysGlnAla 100
QY 301 CTGGACAAGATCGAGAGGAGCAGAAAGTGCAGCAGAAGATCCAGCAGAGCGCGAGGCC 360
Db 101 LeuAspLysIleGluGluGlnGlnAsnLysSerLysLysAlaGlnGlnAla--Ala 119
QY 361 GCCGACAAGGGC-----AAGGTAGCCAGAACTACCCCATCGTGACAGAACTGCAG 411
Db 120 AlaAspThrGlyHisSerAsnGlnValSerGlnAsnTyrProIleValGlnAsnIleGln 139
QY 412 GGCCAGATGCTGCACAGGCCATCAGCCCCCGCACCTGAACGCTGGGTGAAGGTGATC 471
Db 140 GlyIleMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTirpValLysVal 159
QY 472 GAGGAAAGGCGCTTCAGCCCCGAGGTGATCCCATGTTCAACGCGCTGAGCGAGGGGCC 531
Db 160 GluGluLysAlaPheSerProGluValIleProMetPheSerAlaLeuSerGluGlyAla 179
QY 532 ACCCGCCAGGACCTGAACAGATGTTGAACACCGTGGCGGCGCACAGCGCGCATCGAG 591
Db 180 ThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGln 199

QY 592 ATGCTGAAGGACCATCAACGAGAGGCGCGCGAGTGGACCGCGTGACCCCGGTGCAC 651
Db 200 MetLeuLysGluThrIleAsnGluGlnAlaIleGluTirpAspArgValHisProValHis 219
QY 652 GCGGCGCCCATCGCCCCCGGCGAGATGCGCGAGCCCGCGGAGCGACATGCGCGGAC 711
Db 220 AlaGlyProIleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThr 239
QY 712 ACCAGCACCTTCGAGAGCAGATCGCCTGTGATGACCAAGACCCCATCCCGGTGGGC 771
Db 240 ThrSerThrLeuGlnGluGlnIleGlyTyrMetThrAsnAsnProIleProValGly 259
QY 772 GACATCTACAAGCGGTGATCATCTCGGCGCTGAACAAGATCGTGGGATGACAGCCC 831
Db 260 GluIleTyrLysArgTirpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerPro 279
QY 832 GTGAGCATCTGGAACATCAAGCAGGGCCCCCAAGAGCCCTTCGCGACTACGTGACCGC 891
Db 280 ThrSerIleLeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArg 299
QY 892 TTCTTCAAGACCTTCGCGCGCGCGAGCAGAGCACCCAGAGGTGAAGAACTGTGACCGAC 951
Db 300 PheTyrLysThrLeuArgAlaGluGlnAlaSerGlnValLysAsnTirpMetThrGlu 319
QY 952 ACCCTGCTGTGAGAAACGCCAACCCCGACTGCAAGACCATCTCGCGCTCTCGGCCCC 1011
Db 320 ThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyPro 339
QY 1012 GCGCGCAGCCTTGAGAGAGATGATGACCGCCTGCGCAGGGCGTGGCGGCGCCACACA 1071
Db 340 AlaAlaThrLeuGlnGluGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLys 359
QY 1072 GCCCGGTGCTGGCCGCGAGCGATGAGCCAGGCCAACACCAAGC-----GTGATGATGAC 1125
Db 360 AlaArgValLeuAlaGluAlaMetSerGlnValThrAsnSerAlaThrIleMetMetGln 379
QY 1126 AAGACAACCTTCAAGGGCGCCCGCGCATCGTCAAGTGTCTCAACTGCGGCAAGAGGC 1185
Db 380 ArgGlyAsnPheArgAsnGlnArgLysIleValLysCysPheAsnCysGlyLysGluGly 399
QY 1186 CACATCGCCCGCAACTGCCCGCGCCCGCCGCAAGAGGCGTGTGGAAGTGGCGCAAGAG 1245
Db 400 HisThrAlaArgAsnCysArgAlaProArgLysLysGlyCysTirpLysCysGlyLysGlu 419
QY 1246 GGCCACACAGATGAAGAGACTGCACCGAGCGGCGCAACCACTTCTGGGCAAGATCTGGCCC 1305
Db 420 GlyHisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTirpPro 439
QY 1306 AGCCACAAGGGCGCGCCCGGCAACTTCTGAGAGCGGCCCGGAGCCACCGCGCCCGCCC 1365
Db 440 SerTyrLysGlyLysArgProGlyLysPheLeuGlnSerArgProGluProThrAlaProPro 459
QY 1366 GCCGAGAGCTTCGCG-----TTGAGAGAGACACCCCGCGGAGAGAGAGACAAG 1419
Db 460 GluGluSerPheArgSerGlyValGluThrThrThrProProGlnLysGlnGluProIle 479
QY 1420 GACCGCGAG-----ACCCTGACCAAGCTGAAGAGCTGTTCGCAACGAGACCCCTGAGC 1473
Db 480 AspLysGluLeuTyrProLeuThrSerLeuArgSerLeuPheGlyAsnAspProSerSer 499
QY 1474 CAG 1476
Db 500 Gln 500

RESULT 4
US-08-375-510-1
; Sequence 1, Application US/08375510
; Patent No. 5576421
; GENERAL INFORMATION:
; APPLICANT: Saito, Atsushi
; APPLICANT: Sinagawa, Hideo
; APPLICANT: Nakata, Atsuo

```

: TITLE OF INVENTION: HIV ANTIGEN
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Birch, Stewart, Kolasch and Birch
: STREET: P.O. Box 747
: CITY: Falls Church
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22040-0747
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/375,510
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/07/985,949
: FILING DATE: 04-DEC-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Svensson, Leonard R.
: REGISTRATION NUMBER: 30,330
: REFERENCE/DOCKET NUMBER: 216-309P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 241-1300
: TELEFAX: (703) 241-2848
: TELEX: 248345
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 500 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHEetical: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Human immunodeficiency virus type 1
: US-08-375-510-1

Alignment Scores:
Pred. No.: 6.92e-130 Length: 500
Score: 2249.00 Matches: 421
Percent Similarity: 92.22% Conservative: 41
Best Local Similarity: 84.03% Mismatches: 29
Query Match: 81.13% Indels: 10
DB: 1 Gaps: 5

US-09-475-704A-3 (1-1479) x US-08-375-510-1 (1-500)
QY 1 ATGGGGCGCCGCGCCGATCCTGCGCGCGGCAAGCTGACGCGCTGGGAGCGCATCCGC 60
Db 1 MetGlyAlaArgAlaSerValLeuSerGlyGlyGluLeuAspLysTrpGluLysIleArg 20
QY 61 CTGCGCGCGCGCGCGCAAGAGTGCTACATGATGAAGCACTGTGTGGGCCAGCGCGAG 120
Db 21 LeuArgProGlyGlyLysLysGlnTrpLysLeuLysHisIleValTrpAlaSerArgGlu 40
QY 121 CTGAGAGAAGTTCGCGCTGAACCCCGCGCTGTGAGACCAAGGAGGCTGCAAGCAGATC 180
Db 41 LeuGluArgPheAlaValAsnProGlyLeuLeuGluThrSerGluGlyCysArgGlnIle 60
QY 181 ATCCGCCAGCTGCACCCCGCGCTGTCAGACCGGAGGAGGAGCTGAAGAGCTGTTCAAC 240
Db 61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGluGluLeuArgSerLeuTrpAsn 80
QY 241 ACCGTGGCCACCTGTTACTGCGTGTGACAGAGAAGATCGAGGTCCGCGACACCAAGAGAGCC 300
Db 81 ThrIleAlaValLeuTrpCysValHisGlnArgIleAspValLysAspThrLysGluAla 100
QY 301 CTGACAAAGATCGAGAGGAGCAGAAAGAGTGCAGAGAAGATCCAGAGAGCGGAGGCC 360
```

```

Db 101 LeuAspLysIleGluGluGluGlnAsnLysSerLysLysAlaGlnGlnAla---Ala 119
QY 361 GCCGCAAGGCGC-----AAGGTAGCCAGAACTACCCCATGTGCAGAACTGTCAG 411
Db 120 AlaAspThrGlyAsnAsnSerGlnValSerGlnAsnTrpProIleValGlnAsnLeuGln 139
QY 412 GCGCAGATGTGCACCAAGGCATACGCCCGCGACCTGAACGCGCTGGGTGAAGGTGATC 471
Db 140 GlyGlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValVal 159
QY 472 GAGGAGAAAGCCTTACAGCCCGGAGGTGATCCCATGTTACCGCGCTGAGCGAGGCGCC 531
Db 160 GluGluLysAlaPheSerProGluValIleProMetPheSerAlaLeuSerGluGlyAla 179
QY 532 ACCCGCCAGGACCTGAACACGATGTTGAACACCGTGGCGCGCCACAGCGCGCATGCAG 591
Db 180 ThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGln 199
QY 592 ATGCTGAAGACACCATCAACGAGGAGCGCGCGAGTGGAGCCCGGTGCACCCCGTGCAC 651
Db 200 MetLeuLysGluThrIleAsnGluGluAlaAlaGluTrpAspArgLeuHisProValHis 219
QY 652 GCGCGCCCATCGCGCGCGCGCGAGATGCGCGAGCGCGCGCGAGCAGACATGCGCGCAC 711
Db 220 AlaGlyProIleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThr 239
QY 712 ACCAGACCCCTGAGAGAGCAGATGCGCTGTGATGACCAACCCCGCATCCCGTGGGC 771
Db 240 ThrSerThrLeuGlnGluGlnIleGlyTrpMetThrHisAsnProProIleProValGly 259
QY 772 GACATCTACAAGCGGTGATCATCTGTGGCTGGAACAAGATCGTCGATGTACAGCCCC 831
Db 260 GluIleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetLysSerPro 279
QY 832 GTGAGCATCTGACATCAAGCAGAGGCGCCCAAGAGCCCTTCCGCACTACGTGAGCCGC 891
Db 280 ThrSerIleLeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArg 299
QY 892 TTCCTCAAGACCTGCGCGCGCGAGCAGACGACCCAGAGGTGAAGACTGATGACCGAC 951
Db 300 PheTyrLysThrLeuArgAlaGluGlnAlaSerGlnGluValLysAsnTrpMetThrGlu 319
QY 952 ACCCTGTGTGTCAGAACGCCAACCCCGACTGCAAGACCATCTGCGGCTTCGGCCCC 1011
Db 320 ThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyPro 339
QY 1012 GCGCGCAGCCTGAGAGAGATGATGACCGCTGCCAGGCGGTGGCGCGCCAGCCACAAG 1071
Db 340 GlyAlaThrLeuGluGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLys 359
QY 1072 GCCCGGTGCTGGCGGAGCGCATGAGCCAGGCCAAC-----ACAGCGTGAATGTCAG 1125
Db 360 AlaArgValLeuAlaGluAlaMetSerGlnValThrAsnProAlaThrIleMetIleGln 379
QY 1126 AAGAGCAACTTCAAGGGCGCGCGCGCATCGTCAAGTGTCTCAACTGCGGCAAGAGGGC 1185
Db 380 LysGlyAsnPheArgAsnGlnArgLysThrValLysCysPheAsnCysGlyLysGluGly 399
QY 1186 CACATCGCCCGCAACTGCGCGCGCGCGCGCAGAGAAGGCTGCTGGAAGTGGCGCAAGAG 1245
Db 400 HisIleAlaLysAsnCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGlu 419
QY 1246 GGCACCAAGATGAAGGACTGCACCGAGCGCCAGGCCCACTTCTGCGGCAAGATGCGCC 1305
Db 420 GlyHisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpPro 439
QY 1306 AGCCACAAGGGCGCGCGCGGCAACTTCTGCAAGCGCGCGCGCGAGCCACCGCGCGCCC 1365
Db 440 SerHisLysGlyArgProGlyAsnPheLeuGlnSerArgProGluProThrIleProPro 459
QY 1366 GCCGAGAGCTTCGCGCTC-----GAGGAGACCAACCGCGCGCAGAGCAGAGACAAG 1419
```


Db 360 AlaArgValIleuAlaGluAlaMetSerGlnValThrAsnProAlaThrIleMetIleGln 379
QY 1126 AAGAGCAACTTCAGGGCCCCCGCGCATCGTCAAGTCTTCAACTGGCGCAAGAGGGC 1185
Db 380 LysGlyAsnPheArgAsnGlnArgLysThrValLysCysPheAsnGlyLysGluGly 399
QY 1186 CACATCGCCCGCAACTGCGCGCGCCCCCGCAAGAGGCTGTGAAGTGGCGCAAGGAG 1245
Db 400 HisIleAlaLysAsnGlySerGlnAlaProArgLysLysGlyCysTrpLysCysGlyLysGlu 419
QY 1246 GGCCACGATGAAGAGTGCACCGAGCGCCAGGCCAACTTCTGGGCAAGATCTGGCC 1305
Db 420 GlyHisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpPro 439
QY 1306 AGCCACAGGGCGCGCGCGCAACTTCTGCAGAGCGCGCGCCAGCCGCGCGCGCGCC 1365
Db 440 SerHisLysGlyArgProGlyAsnPheLeuGlnSerArgProGluProThrAlaProPro 459
QY 1366 GCCGAGACTTCGCTTC-----GAGGAGACCAACCGCGCGCGCAAGAGAGCAAG 1419
Db 460 GluGlnSerPheArgPheGlyGluGlnThrThrProSerGlnLysGlnLysProIle 479
QY 1420 GACCGCGAG-----ACCCTGACGAGCTGAAGAGCTGTTCGCAAGACCCCTGAGC 1473
Db 480 AspLysGluLeuTyrProLeuAlaSerLeuArgSerLeuPheGlySerAspProSerSer 499
QY 1474 CAG 1476
Db 500 Gln 500

RESULT 6
US-09-309-572-16
; Sequence 16, Application US/09309572
; Patent No. 6440730
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral hybrid vectors pseudotyped with LCMV
; FILE REFERENCE: P50489
; CURRENT APPLICATION NUMBER: US/09/309,572
; CURRENT FILING DATE: 1999-05-11
; EARLIER APPLICATION NUMBER: DE 198 56 463
; EARLIER FILING DATE: 1998-11-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: gag polypeptide
US-09-309-572-16

Alignment Scores:
Pred. No.: 6.92e-130 Length: 500
Score: 2249.00 Matches: 421
Percent Similarity: 92.22% Conservative: 41
Best Local Similarity: 84.03% Mismatches: 29
Query Match: 81.13% Indels: 10
DB: 4 Gaps: 5

US-09-475-704A-3 (1-1479) x US-09-309-572-16 (1-500)

QY 1 ATGGGCGCGCGCGCGCATCTCGCGCGCGCGCAAGCTGGAGCGCTGGGAGCGCATCCGC 60
Db 1 MetGlyAlaArgAlaSerValLeuSerGlyGlyLeuAspLysTrpGluLysIleArg 20
QY 61 CTGCGCGCGCGCGCGCAAGAGTGTACATGATGAAGCACCTGTGTGGGCCAGCGCGAG 120
Db 21 LeuArgProGlyGlyLysLysGlnTyrLysLeuLysHisIleValItrpAlaSerArgGlu 40
QY 121 CTGAGAGAGTTGCGCCCTGAACCCCGCGCTGTGAGACCAAGCGGCTGCAAGCAGATC 180
Db 41 LeuGluArgPheAlaValAsnProGlyLeuLeuGluThrSerGlnGlyCysArgGlnIle 60

QY 181 ATCCGCCAGCTGCAACCCCGCTGACAGACCGGAGGAGAGCTGAAGAGCTGTTCAC 240
Db 61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGluGluLeuArgSerLeuTyrAsn 80
QY 241 ACCGTGGCCACCTGTACTGCGTGCACAGAGAAGATCGAGTCCCGCACACCAAGAGGCC 300
Db 81 ThrIleAlaValLeuTyrCysValHisGlnArgIleAspValLysAspThrLysGluAla 100
QY 301 CTGGAACAAGATCGAGAGAGAGCAACAAGTCCAGACAAGATCCAGAGCGCGAGGCC 360
Db 101 LeuAspLysIleGluGluGluGlnAsnLysSerLysLysLysAlaGlnGlnAla--Ala 119
QY 361 GCCGACAAGGC-----AAGGTGAGCCAGAACTACCCCATCTGTGCAGAAGCTGCAG 411
Db 120 AlaAspThrGlyAsnAsnSerGlnValSerGlnAsnTyrProIleValGlnAsnLeuGln 139
QY 412 GGCCAGATGTGTCAACGAGCCATCAGCCCGCGCACCTGAACGCTGGGTGAAGGTGATC 471
Db 140 GlyGlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaItrpValLysValVal 159
QY 472 GAGAGAAAGGCTTCAGCCCGCGAGTGTATCCCATGTTCAACCGCGCTGAGCGAGGGCC 531
Db 160 GluGluLysAlaPheSerProGluValIleProMetPheSerAlaLeuSerGluGlyAla 179
QY 532 ACCCGCCAGGAGCTGAACAACGATGTGAACACCGGTGGCGGCGCACAGCGCGCATGCA 591
Db 180 ThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaIleMetGln 199
QY 592 ATGCTGAAGAGACCATCAACGAGAGAGCGCGCGAGTGGAGCCGCTGCACCCCGTGAC 651
Db 200 MetLeuLysGluThrIleAsnGluGluAlaAlaGluTrpAspArgLeuHisProValHis 219
QY 652 GCGGCGCCCATCGCGCGCGCGAGATGCGCGAGCGCGCGCGAGCATCGCGGCGCAC 711
Db 220 AlaGlyProIleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThr 239
QY 712 ACCAGCACCTTGAGAGAGAGCATGCGCTGTGATGACCAAGACCCCGCATCCCGTGGC 771
Db 240 ThrSerThrLeuGlnGluGlnIleGlyTyrMetThrHisAsnProProIleProValGly 259
QY 772 GACATCTACAAGCGGTGATCATCTGCGCTGAACAAGATCGTGGATGTACAGCCCC 831
Db 260 GluIleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerPro 279
QY 832 GTGAGCATCTTGACATCAAGCAGAGCGCGCGCAAGCGCTTCGCGACTACGTGAGCCGC 891
Db 280 ThrSerIleLeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArg 299
QY 892 TTCTTCAAGACCTGCGCGCGCGAGCAGAGCACCCAGAGGTGAAGTGAAGTGAACCGAC 951
Db 300 PheTyrLysThrLeuArgAlaGluGlnAlaSerGlnGluValLysAsnTyrMetThrGlu 319
QY 952 ACCGTGCTGTGAGAGAGCAAGCGCAACCCCGACTGCAAGCATCTGCGGCTTGGCCCC 1011
Db 320 ThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyPro 339
QY 1012 GGCGCGAGCTTGAGAGAGATGATGACCGCTGCCAGGCGGTGGCGCGCGCGCAAG 1071
Db 340 GlyAlaThrLeuGluGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLys 359
QY 1072 GCCCGGTGTGCGCGCGAGCGAGTGAAGCGCAAC-----ACCAGCGTGAATGATGCAG 1125
Db 360 AlaArgValLeuAlaGluAlaMetSerGlnValThrAsnProAlaThrIleMetIleGln 379
QY 1126 AAGAGCAACTTCAAGGGCGCGCGCGCATCGTCAAGTGTTCAACTGGCGCAAGAGGC 1185
Db 380 LysGlyAsnPheArgAsnGlnArgLysThrValLysCysPheAsnGlyLysGluGly 399
QY 1186 CACATGCGCGCGCAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1245
Db 400 HisIleAlaLysAsnGlySerGlnAlaProArgLysLysGlyCysTrpLysCysGlyLysGlu 419

QY	1246	GGCCAC	CAGATGA	AGGACT	GCAC	CCGAGCGCC	AGCCAA	CTTCTCTG	GGCAAGATCT	GGCCC	1303	
Db	420	GlyHis	GlnMet	LeuLys	AspCys	ThrGlu	ArgGln	AlaAsn	PheLeu	GlyLys	IleTrp	Pro 439
QY	1306	AGCCACA	AGGGCG	CGCCCC	CGGCA	CTTCTCTG	CAGACCG	CGCCCC	CGGACCC	CGCCCC	1365	
Db	440	SerHis	LeuSer	GlyArg	ProGlu	AsnPhe	LeuGln	SerArg	ProGlu	ProThr	AlaPro	Pro 459
QY	1366	GCCGAG	AGCTTCC	GCCTTC	-----	GAGGAG	ACCA	CCCCCG	CCGACA	AGCAGAG	AGCAAG	1419
Db	460	GluGln	SerPhe	ArgPhe	GlyGlu	GluThr	ThrThr	ProSer	GlnLys	GlnGlu	ProIle	479
QY	1420	GACCGC	GAG-----	ACCCTG	ACCA	GCCTGA	AGAG	CCTGTT	GGCA	AGCACC	CCCTGAGC	1473
			:::			:::		:::		:::		
Db	480	AspLys	GluLeu	TyrPhe	LeuAla	SerLeu	ArgSer	LeuPhe	GlySer	AspPro	SerSer	499
QY	1474	CAG	1476									
Db	500	Gln	500									

```

RESULT 7
US-09-718-096-16
; Sequence 16, Application US/09718096
; Patent No. 6589763
; GENERAL INFORMATION:
; APPLICANT: Von Laer, Meike-Dorothee
; TITLE OF INVENTION: RETROVIRAL HYBRID VECTORS PSEUDOTYPED WITH LCMV
; FILE REFERENCE: 35-195
; CURRENT APPLICATION NUMBER: US/09/718,096
; CURRENT FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: DE 19856463.5
; PRIOR FILING DATE: 1998-11-26
; PRIOR APPLICATION NUMBER: EP 99250415.9
; PRIOR FILING DATE: 1999-11-25
; PRIOR APPLICATION NUMBER: US 09/309,572
; PRIOR FILING DATE: 1999-05-11
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: gag polyprotein
US-09-718-096-16

```

Alignment Scores:	
Pred. No.:	6.92e-130
Score:	2249.00
Percent Similarity:	92.22%
Best Local Similarity:	84.03%
Query Match:	81.13%
DB:	4
Length:	500
Matches:	421
Conservative:	41
Mismatches:	29
Indels:	10
Gaps:	5

US-09-475-704A-3 (1-1479) X US-09-718-096-16 (1-500)

```

QY      1 ATGGGCGCCCGCGCCAGCATCCCTGCGCGGCGGCAAGCTGGACCGCCTGGAGCGCATCCG  60
Db      1 MetGlyAlaArgAlaSerValLeuSerGlyGlyLeuAspLysTrpGluLysIleArg  20
QY      61 CTGGCGCCCGCGCGGCAAGAGTGTCTACATGATGAAGCACCCTGTGTGGCGCAGCCGCGAG  120
Db      21 LeuArgProGlyGlyLysLysGlnTyrLysLeuLysHisIleValTrpAlaSerArgGlu  40
QY      121 CTGGAGAAGTTCGCCCTGTAACCCCGCGCTGCTGGAGACCGAGGCGCTGCAAGCAGATC  180
Db      41 LeuGluArgPheAlaValAsnProGlyLeuLeuGlnThrSerGlnGluLysArgGlnIle  60
QY      181 ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGCGAGCGAGAGCTGAGACGCTGTTCAAC  240
Db      61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGlnGluLeuArgSerLeuTyrAsn  80
QY      241 ACCGTGGCCACCCCTGTACTGCGTGCAGCAGAAGATCGAGTCCGCGACACCAAGAGGCC  300

```

Db	81	ThrlleAlaValleuTyrCysValHisGlnArgIleAspValIlyAspThrIySGluAla	100
QY	301	CTGGACAAGATCGAGAGAGACAAGTGCACAGAGATCCAGCGCGAGCC	360
Db	101	LeuAspIySIlleGIuGIuGIuAsnIySserIySbIySAlaGInGlnAla--Ala	119
QY	361	GCCGACAAGGC-----AAGTGAGCCAGACTACCCCATCTGCAGACCTGCAG	411
Db	120	AlaAspThrGIyAsnAsnSerGlnValSerGlnAsnTyrProIleValGlnAsnLeuGln	139
QY	412	GGCCAGATGTGCAACCAGGCATACCCCCGACCCCTGAACGCCCTGGTGAAGGTGATC	471
Db	140	GIyGlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaIrrValIlySValVal	159
QY	472	GAGGAGAAGGCCCTTCAGCCCCGAGGTGATCCCATGTTCACCGCCCTTGAAGGGCGCC	531
Db	160	GIuGIuIySAlaPheSerProGIuValIleProMetPheSerAlaIleuSerGIuGIyAla	179

QY	532	ACCCCCCAGGACCTGAAACAGCA	TGTTGAACACCCGTGGGGGCCAC	CGAGGCCGCGCATGCA	591
Db	180	ThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlnHisGlnAlaMetGln			199
QY	592	ATGCTGAAGGACACCATCAACGAGAGGGCCGCGAGTGGACCGCGTGCACCCCGTGAC			651
Db	200	MetLeuLysGlnThrIleAsnGlnGluAlaGluTrpAspArgLeuHisProValHis			219
QY	652	GGCGGGCCCATCGCCCCCGGCGAGATGCGCGGAGCCCCCGGCGAGCATCGCGGAC			711
Db	220	AlaGlyProIleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThr			239
QY	712	ACCAGCACCCCTGCAGAGAGACAGATCGCCTGGATGACCAACCCCCCATCCCGTGGG			771
Db	240	ThrSerThrLeuGlnGlnGlnIleGlyTrpMetThrHisAsnProProIleProValGly			259
QY	772	GACATCTACAACGGGTGATCATCTGGGCGCTGAACAAGATCGTGGGATPACAGCCC			831
Db	260	GluIleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerPro			279
QY	832	GTGAGCATCCTGGAACATCAACGAGGGCCCGAAGAGCCCTCCGCACTAGCTGACCG			891
Db	280	ThrSerIleLeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArg			299
QY	892	TTCTTCAAGACCTGCGCGCGCCGAGCAGAGCAACCCAGAGGTGAAGAACTGTAACCG			951
Db	300	PheTyrLysThrLeuArgAlaGlnAlaSerGlnGluValLysAsnTrpMetThrGlu			319
QY	952	ACCTGTGTGTGACAGAACGCCAACCCCGCATGCAAGACCATCTGCGGCTCTGGCCCC			1011
Db	320	ThrLeuLeuValGlnAsnAlaAsnProAspLysThrIleLeuLysAlaLeuGlyPro			339
QY	1012	GGCGCCAGCCTGAGAGAGATGATGACCGGCTGCGAGGGCGTGGGCGGCCAGCCACAG			1071
Db	340	GlyAlaThrLeuGlnGluMetMetThrAlaCysGlnGlyValGlyLyrProGlnHisLys			359
QY	1072	GGCCGCGTGTGGCCGAGGCGCATGAGCCAGGCCAAC-----ACCAGCGTATGATGCAG			1125
Db	360	AlaArgValLeuAlaGluAlaMetSerGlnValThrAsnProAlaThrIleMetIleGln			379
QY	1126	AAGAGCAACTTCAAGGGGCCCCCGCGCATGTCATAGTCTTCAACTGCGGCAAGAGGG			1185
Db	380	LysGlnAsnPheArgAsnGlnArgLysThrValLysCysPheAsnCysGlyLysGlnGly			399
QY	1186	CACATCGCCCGCACTGCGCGGCCCCCGGCAAGAGGGCTGTGAAGTGCGGCAAGAG			1245
Db	400	HisIleAlaLysAsnCysAlaGlnAlaProArgLysLysGlyCysTrpLysCysGlyLysGln			419
QY	1246	GGCCACCAAGATGAAGACTGCAACGAGCGCCAGGCCAACTTCTCTGGGCAAGATCTGG			1305
Db	420	GlyHisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpPro			439
QY	1306	AGCCACAAAGGGCGCGCCCGGCACTTCTGTGAGAGCGCGCCGAGCCACCGCGCCCC			1365

Db 440 SerHisIySGlyArgProGlyAsnPheLeuGlnSerArgProGluProThrAlaProPro 459
QY 1366 GCCGAGAGCTTCGGCTTC-----GAGAGACCAACCCCGCCAGAACGAGAGAGCAAG 1419
Db 460 GluGlnSerPheArgPheGlyGluGlnThrThrProSerGlnIleGlnGluProIle 479
QY 1420 GACCGCGAG-----ACCCTGACGAGCTGGAAGAGCTGTTCGCAACGACCCCTGAGC 1473
Db 480 AspIySGluLeuTyrProLeuAlaSerLeuArgSerLeuPheGlySerAspProSerSer 499
QY 1474 CAG 1476
Db 500 Gln 500
RESULT 8
US-09-952-060-35
; Sequence 35, Application US/09952060
; Patent No. 6733993
; GENERAL INFORMATION:
; APPLICANT: Emini, Emilio A.
; APPLICANT: Youil, Rima
; APPLICANT: Bett, Andrew J.
; APPLICANT: Chen, Ling
; APPLICANT: Kaslow, David C.
; APPLICANT: Shiver, John W.
; APPLICANT: Toner, Timothy J.
; APPLICANT: Casimiro, Danilo R.
; TITLE OF INVENTION: ENHANCED FIRST GENERATION ADENOVIRUS
; TITLE OF INVENTION: VACCINES EXPRESSING CODON OPTIMIZED HIV1-GAG, POL, NEF AND
; TITLE OF INVENTION: MODIFICATIONS
; FILE REFERENCE: 20747Y
; CURRENT APPLICATION NUMBER: US/09/952, 060
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: PCT/US01/28861
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/317, 814
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/279, 056
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/233, 180
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 1350
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Codon optimized gag-1A pol fusion
US-09-952-060-35
Alignment Scores:
Pred. No.: 1,44e-129 Length: 1350
Score: 2245.00 Matches: 419
Percent Similarity: 92.00% Conservative: 41
Best Local Similarity: 83.80% Mismatches: 32
Query Match: 80.99% Indels: 8
DB: 4 Gaps: 4
US-09-475-704A-3 (1-1479) x US-09-952-060-35 (1-1350)
QY 1 ATGGCGCCCGCGCCAGCATCTGCGCGCGGCAAGCTGAGCGCTGGGAGCGCATCCGC 60
Db 1 MetGlyAlaArgAlaSerValLeuSerGlyGluLeuAspLysTrpGluLysIleArg 20
QY 61 CTGCGCCCGCGCGCAAGAGTGCTACATGATGAAGCACTGTGTGGCCAGCGCGAG 120
Db 21 LeuArgProGlyGlyLysLysLysTyrLysLeuLysHisIleValTrpAlaSerArgGlu 40
QY 121 CTGAGAGATTGCGCTGAAACCCCGCGCTGTGAGACGAGCGGCTGCAAGCAGATC 180
Db 41 LeuGluArgPheAlaValAsnProGlyLeuLeuGluThrSerGluGlyCysArgGlnIle 60

QY 181 ATCCGCCAGCTGACCCCGCTGACAGCCGAGCGAGAGCTGAAGCTGTTCAC 240
Db 61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGluGluLeuArgSerLeuTyrAsn 80
QY 241 ACCGTGGCACCTGTACTGCGTGACGAGAAGATCGAGTCCGCGACCAAGAGGCC 300
Db 81 ThrValAlaThrLeuTyrCysValHisGlnLysIleAspValLysAspThrLysGluAla 100
QY 301 CTGACAAGATCGAGAGAGAGCAACAAGTGCCAGACAGAAGATCCAGACGCCGAGGCC 360
Db 101 LeuGluLysIleGluGluGluGlnAsnLysSerLysLysAlaGlnGlnAlaAla 120
QY 361 -----GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGACAGACCTGACGGCC 414
Db 121 GlyThrGlyAsnSerSerGlnValSerGlnAsnTyrProIleValGlnAsnLeuGlnGly 140
QY 415 CAGATGCTGACACGAGCCCATCAGCCCGCCGACCCCTGAACGCTGGGTGAAGTGAATCGAG 474
Db 141 GlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValGlu 160
QY 475 GAGAGGCTTCAAGCCCGAGGTATCCCATGTTCAACGCGCTGACGAGGCGCCACC 534
Db 161 GluLysAlaPheSerProGluValIleProMetPheSerAlaLeuSerGluGlyAlaThr 180
QY 535 CCCCAAGACCTGAACACGATGTTGAACACCGTGGCGGCCACACGCGCCCATGAGATG 594
Db 181 ProGlnAspLeuAsnThrMetLeuAsnThrValGlyHisGlnAlaIleMetGlnMet 200
QY 595 CTGAAGACACCATCAACGAGAGAGCGCCGAGTGGGACCGCGCTGACACCCCGTGACGCC 654
Db 201 LeuLysGluThrIleAsnGluGlnAlaIleGluTrpAspArgLeuHisProValHisAla 220
QY 655 GGGCCCATCGCCCCCGCCAGATGCGGAGAGCCCGCGGACGACATCGCCGACCAACC 714
Db 221 GlyProIleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThr 240
QY 715 AGCACCTTGACAGAGCAAGATCGCTTGATGACCAACACCCCATCCCGTGGCGGAC 774
Db 241 SerThrLeuGlnGlnIleGlyTyrMetThrAsnAsnProIleProValGlyGlu 260
QY 775 ATCTAACAGCGGTGATCATCTCTGGGCTGAACAAGATCGTGGGATGTACAGCCCGTG 834
Db 261 IleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProThr 280
QY 835 AGCATCTTGACATCAAGCAGGGGCCCCAAGGAGGCTTCCGCGACTACGTGACGCTTC 894
Db 281 SerIleLeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhe 300
QY 895 TTCAAGACCTTGGCGCGCCGAGAGACGACCCAGAGAGGTGAAGAATGTGATGACGACACC 954
Db 301 TyrLysThrLeuArgAlaGluGlnAlaSerGlnGlnValLysAsnTrpMetThrGluThr 320
QY 955 CTGCTGTGACAGAACGCCAACCCCGACTGCAAGACCATCTTGCGCGCTCTCGCGCCGCGC 1014
Db 321 LeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyProAla 340
QY 1015 GCCAGCTTGAGAGATGATGATGACCGCTGCCAGGCGGTGGCGGCCCCAGCCACAAGGCC 1074
Db 341 AlaThrLeuGlnGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLysAla 360
QY 1075 CCGGTGCTGGCGAGGGCGATGAGCCAGCCAAACCCAGC-----GTGATGATGACAGAAG 1128
Db 361 ArgValLeuAlaGluAlaMetSerGlnValThrAsnSerAlaThrIleMetMetGlnArg 380
QY 1129 AGCACTTCAAGGAGCCCCCGCGCATGCTCAAGTCTTCAATCGCGGCAAGAGGCGCAC 1188
Db 381 GlyAsnPheArgAsnGlnArgLysThrValLysCysPheAsnLysGlyLysValGlyHis 400
QY 1189 ATGCGCCGCAACTGCGCGCGCCCGCCGCAAGAGGGCTGTGAAGTGCGGCAAGAGGGC 1248
Db 401 IleAlaLysAsnLysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGluGly 420
QY 1249 CACCAATGAAGACTGACACCGAGCGCCAGGCGCAACTTCTGTGGCAAGATCTGGCCAGC 1308

```

Db      421 HisGlnMetLysAspCysAsnGluArgGlnAlaAsnPhenLeuGlySerIleTrpProSer 440
QY      1309 CACAAAGGCGCGCCCGGCAACTTCTGTGACAGAGCGCGCCGAGCCCAAGCCCGCCCGCCGCC 1368
Db      441 HisIysGlyArgProGlyAsnPhenLeuGlnSerArgProGluProThrAlaProProGlu 460
QY      1369 GAGAGCTCCGCTTC-----GAGAGACCAACCCCGCCAGAACGACGAGAGACGAGAC 1422
Db      461 GluSerPheArgPheGlyGlyGluIlyThrThrProSerGlnLysGlnGluProIleAsp 480
QY      1423 CGCGAG-----ACCTGACCAGCCTGAAGAGCCTGTTGGCAACGACCCCTGAGCCAG 1476
Db      481 LysGluLeuTyrProIleuAlaSerLeuArgSerLeuPheGlyAsnAspProSerSerGln 500

```

RESULT 9

```

US-08-463-210-8
; Sequence 8, Application US/08463210
; Patent No. 6001977
; GENERAL INFORMATION:
; APPLICANT: CHANG, Nancy T.
; APPLICANT: GALLO, Robert C.
; APPLICANT: WONG-STAL, Floesle
; TITLE OF INVENTION: CLONING AND EXPRESSION OF HTLV-III DNA
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,210
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/693,866
; FILING DATE: 23-JAN-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/659,339
; FILING DATE: 10-OCT-1984
; ATTORNEY/AGENT INFORMATION:
; NAME: Serunian, Leslie A.
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-4193US2
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 512 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: HTLV-III
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..512
; OTHER INFORMATION: /note= "gag protein of HTLV-III"
US-08-463-210-8

```

Alignment Scores:

```

Pred. No.: 4.99e-129 Length: 512
Score: 2235.00 Matches: 424

```

```

Percent Similarity: 90.64% Conservative: 41
Best Local Similarity: 82.65% Mismatches: 26
Query Match: 80.63% Indels: 22
DB: 3 Gaps: 7

```

US-09-475-704A-3 (1-1479) x US-08-463-210-8 (1-512)

```

QY      1 ATGGGCGCCCGCCAGCATCTGCGCGCGCAAGCTGAGCGCTGGAGCCATCCGC 60
Db      1 MetGlyAlaArgAlaSerValLeuSerGlyGlyGluLeuAspArgTrpGluSerIleArg 20
QY      61 CTGCGCCCGCGCGCAAGAGTGTCTACATGATGAAGCAACCTGTGTGGGCCAGCCGAG 120
Db      21 LeuArgProGlyGlyLysLysTyrTrpLysLeuLysHisIleValTrpAlaSerArgGlu 40
QY      121 CTGAGAGAGTTGCGCCCTGAACCCCGCGCTGTGAGACCAAGCGAGGCTGCAAGCATC 180
Db      41 LeuGluArgPheAlaValAlaAsnProGlyLeuLeuGluThrSerGlnGlyCysArgGlnIle 60
QY      181 ATCCGCGAGCTGCAACCCCGCGCTGACAGCCGCGACGAGAGCTGAAGAGCTGTTCAC 240
Db      61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGlnGluLeuArgSerLeuTyrAsn 80
QY      241 ACCGTGGCCACCTGTACTGCTGTCACGAGAGATCGAGGTCCGCGACCAAGAGGCC 300
Db      81 ThrValAlaThrLeuTyrCysValHisGlnArgIleGluIleLysAspThrLysGluAla 100
QY      301 CTGAGACAAGATGAGAGAGAGAGACAACAGTGCACGAGAAATCCAGACGCGGAGGCC 360
Db      101 LeuAspLysIleGluGluGluGlnAsnLysSerLysLysValAlaGlnGlnAla--Ala 119
QY      361 GCCGACAAGGCG-----AAGGTGAGCCAGAACTACCCCATCGTGACAACTGACAG 411
Db      120 AlaAspThrGlyHisSerSerGlnValSerGlnAsnTyrProIleValGlnAsnIleGln 139
QY      412 GGCACAGATGTCACACAGGCCATCAGCCCGCCACCCCTGAAGCCTGGGTGAAGTGATC 471
Db      140 GlyGlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValVal 159
QY      472 GAGGAGAAGGCTTCAGCCCGGAGGTGATCCCATGTTCAACGCGCCCTGAGCGAGCGGCC 531
Db      160 GluGluLysAlaPheSerProGluValIleProMetPheSerAlaLeuSerGlnGlyAla 179
QY      532 ACCCCCGAGACTGAACACAGATGTTGAACCGTGGCGGCGCACCGCCCATGACAG 591
Db      180 ThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlnHisGlnAlaIleMetGln 199
QY      592 ATGCTGAAGACCATCAACGAGAGAGCGCGCGAGTGAGACCGCGTGACCCCGTGAC 651
Db      200 MetLeuLysGluThrIleAsnGluGluAlaIleGluTrpAspArgValHisIleProValHis 219
QY      652 GCCGCGCCCATGCGCCCGCGGAGATGCGCGAGCCCGCGGAGCGACATCGCCGCGCAC 711
Db      220 AlaGlyProIleAlaProGlyGlnMetArgGluProArgLysSerAspIleIleGlyThr 239
QY      712 ACCAGCACCTTGAGAGAGAGATCGCTGATGACCAACCCCGCATCCCGGTGGGC 771
Db      240 ThrSerThrLeuGlnGluGlnIleGlyTrpMetThrAsnAsnProProIleProValGly 259
QY      772 GACATCTACAAGCGGTGATCATCTGTGGCTGAACAAGATGTCGATGACAGCCCG 831
Db      260 GluIleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetCysSerPro 279
QY      832 GTGAGCATCTTGACATCAAGCAGAGCGCCCAAGAGAGCCCTTCGCGGACTAGTGAGCCGC 891
Db      280 ThrSerIleLeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArg 299
QY      892 TTCTTCAAGACCTTGCGCGCGCGAGCAGACCAAGAGGTGAAGAACTGATGACCGCAC 951
Db      300 PheTyrLysThrLeuArgAlaGluGlnAlaSerGlnGluValLysAsnTrpMetThrGlu 319
QY      952 ACCCTGTGTGAGAGAGCGCAACCCCGACTGCAAGACCATCTGCGCGCTTCGCGCCCC 1011

```

Db 320 ThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyPro 339
QY 1012 GGGCCACGCTTGAGAGATGATGACCCGCTGCCAGGGCGGTGGGGCCCCCAGCCACAG 1071
Db 340 AlaAlaThrLeuGlnGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLys 359
QY 1072 GCGCGCGTCTGGCCGAGCGGATGAGCCAG--GCCAACACC--AGCGTATGATGAG 1125
Db 360 AlaArgValLeuAlaGlnAlaMetSerGlnValThrAsnThrAlaThrIleMetMetGln 379
QY 1126 AAGAGCACTCAAGGGGCCCCCGCGCATGCTCAAGTGTCTCAACTGGCGCAAGAGGGC 1185
Db 380 ArgGlyAsnPheArgAsnGlnArgLysMetValLysCysPheAsnGlyLysGlnGly 399
QY 1186 CACATCGCCCGCAACTGCGCGCGCCCCCGCAAGAGGGCTGCTGAAAGTGGCGCAAGAG 1245
Db 400 HisthrAlaArgAsnCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGln 419
QY 1246 GGCACACAGATGAAGACTGCACCCGAGCGCCAGCCAACTTCCTGGCAAGATCTGGCCC 1305
Db 420 GlyHisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpPro 439
QY 1306 AGCCACAAGGGCGCGCGCGCAAC----- 1329
Db 440 SerTyrLysGlyArgProGlyAsnPheLeuGlnSerArgProGluProThrAlaProPro 459
QY 1330 TTCCTGCAGAGCGCGCGCGAGCCGCGCCCGCGCGAGAGCTTCGCG-----TTC 1383
Db 460 PheLeuGlnSerArgProGluProThrAlaProProGluGlnSerPheArgSerGlyVal 479
QY 1384 GAGGAGACCAACCCCGCGCAG 1437
Db 480 GluThrThrThrProProGlnLysGlnGluProIleAspLysGluLeuTyrProLeuThr 499
QY 1438 AGCTGAAGAGCCTGTTGCGCAACGAGCCCTGAGCCAG 1476
Db 500 SerLeuArgSerLeuPheGlyAsnAspProSerSerGln 512
RESULT 10
US-09-124-900-2
; Sequence 2, Application US/09124900
; Patent No. 6268484
; GENERAL INFORMATION:
; APPLICANT: KATINGER, Hermann
; APPLICANT: BUCHACHER, Andrea
; APPLICANT: ERNST, Wolfgang
; APPLICANT: BALAUN, Claudia
; APPLICANT: PUTSCHER, Martin
; APPLICANT: TRKOLA, Alexandra
; APPLICANT: PREDL, Renate
; APPLICANT: SCHMATZ, Christine
; APPLICANT: KLIMA, Annelies
; APPLICANT: STEINDL, Franz
; APPLICANT: MUSTER, Thomas
; TITLE OF INVENTION: HIV-Vaccines
; FILE REFERENCE: 1939-112P
; CURRENT APPLICATION NUMBER: US/09/124, 900
; CURRENT FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: PCT/EP95/01481
; PRIOR FILING DATE: 1995-04-19
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-124-900-2

Query Match: 80.63% Indels: 22
DB: 3 Gaps: 7
US-09-475-704A-3 (1-1479) x US-09-124-900-2 (1-512)
QY 1 ATGGCGCGCGCGCGCAGCATCTCGCGCGCGCGCGCAAGCTGACCGCCTGGAGCGCATCCGC 60
Db 1 MetGlyAlaArgAlaSerValLeuSerGlyGlyGluLeuAspArgTrpGluLysIleArg 20
QY 61 CTGCGCGCGCGCGCGCAAGAAGTGTACATGATGAGACACCTGTGTGGCCAGCGCGAG 120
Db 21 LeuArgProGlyGlyLysLysLysTyrLysLeuLysHisIleValTrpAlaSerArgGlu 40
QY 121 CTGAGAAAGTTCGCGCTGAACCGCGCGCTGTGAGACACGAGCGGCTGCAAGAGATC 180
Db 41 LeuGluArgPheAlaValAsnProGlyLeuLeuGluThrSerGluGlyCysArgGlnIle 60
QY 181 ATCCGCCAGCTGCACCCCGCGCTGCAGACCGCGCAGCGAGAGCTGAAGAGCCTGTTCAC 240
Db 61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGluGluLeuArgSerLeuTyrAsn 80
QY 241 ACCGTGGCCACCCCTGTACTGCGGTGCACAGAGAGATCGAGGTCCGCGCAGACCAAGAGGCC 300
Db 81 ThrValAlaThrLeuTyrCysValHisGlnArgIleGluIleLysAspThrLysGluAla 100
QY 301 CTGAGACAAGATCGAGAGAGAGAGAGCAACAAGTGCACAGAGATCCAGACGCGCGAGGCC 360
Db 101 LeuAspLysIleGluGluGluGlnAsnLysSerLysLysAlaGlnGlnAla--Ala 119
QY 361 GCCGACAAGGCG-----AAGGTGAGCCAGAACTACCCCATCTGTGCAGAACTGCAG 411
Db 120 AlaAspThrGlyHisSerSerGlnValSerGlnAsnTyrProIleValGlnAsnIleGln 139
QY 412 GGCAGATGTGTACAGGCCCATCAGCCCGCGCACCCTGAACGCGCTGGTGAAGTGATC 471
Db 140 GlyGlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValVal 159
QY 472 GAGGAGAAAGGCTTCAGCCCGCGAGGTGATCCCATGTTCAACGCGCCTGAGCGAGCGCGCC 531
Db 160 GluGluLysAlaPheSerProGluValIleProMetPheSerAlaLeuSerGluGlyAla 179
QY 532 ACCCGCCAGGACCTGAACAGATGTGAACACCGTGGCGCGCGCAGCGCCGATGCAG 591
Db 180 ThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGln 199
QY 592 ATGCTGAAGACACCATCAACGAGAGAGCGCGCGAGTGGACCGGTGCACCCGCTGCAC 651
Db 200 MetLeuLysGluThrIleAsnGluGluAlaIaIaGluTrpAspArgValHisProValHis 219
QY 652 GCGCGCGCGCATGCGCGCGCGCAGATGCGCGAGCCCGCGCGCAGACATCGCGCGCACC 711
Db 220 AlaGlyProIleAlaProGlyGlnMetArgGluProArgLysSerAspIleAlaGlyThr 239
QY 712 ACCAGACCCCTGCAGAGAGAGATGCGCTGATGACCAAGACCCCGCATCCCGTGGGC 771
Db 240 ThrSerThrLeuGlnGluGlnIleGlyTyrMetThrAsnAsnProProIleProValGly 259
QY 772 GACATCTACAAGCGGTGATGATCTCTGGGCTGAACAAGATGTCGATGTACAGCCCG 831
Db 260 GluIleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetCysSerPro 279
QY 832 GTGAGCATCTTGACATCAAGCAGGGCGCCCAAGAGAGCCCTTCGCGACTACGTGAGCGC 891
Db 280 ThrSerIleLeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArg 299
QY 892 TTCTTCAAGACCTTGCGCGCGCAGAGAGACCAAGAGGTGAAGACTGTGACCGCAC 951
Db 300 PheTyrLysThrLeuArgAlaGluGlnAlaSerGlnGluValLysAsnTrpMetThrGlu 319
QY 952 ACCGTGCTGTGCAGAAAGCGCAACCCGAGTGCAGAACCATCTCGCGCGCTCGGCGCCC 1011
Db 320 ThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyPro 339

Alignment Scores:
Pred. No.: 4.99e-129 Length: 512
Score: 2235.00 Matches: 424
Percent Similarity: 90.64% Conservative: 41
Best Local Similarity: 82.65% Mismatches: 26

QY 1012 GCGCCAGCCTGAGAGATGATGACCCCTGCCAGGGCGTGGCGGCCCCAGCCACAAG 1071
Db 340 AlaAlaThrLeuGluGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLys 359
QY 1072 GCGCGGTGCTGGCCGAGCGATGAGCCAG--GCCACAACC--AGCGTATGATGACAG 1125
Db 360 AlaArgValLeuAlaGluAlaMetSerGlnValThrAsnThrAlaThrIleMetGln 379
QY 1126 AAGAGCACTTCAAGGGCCCCCGCGCATCGTCAAGTCTTCAACTGCGGCAAGAGGCC 1185
Db 380 ArgGlyAsnPheArgAsnGlnArgLysMetValLysCysPheAsnGlyLysGlu 399
QY 1186 CACATCGCCCGCAACTGCGCGCGCGCGCGCGCAAGAGGCTGCTGAAGTGGCAAGAG 1245
Db 400 HisThrAlaArgAsnCysAlaProArgLysLysGlyCysTrpLysCysGlyLysGlu 419
QY 1246 GCGCCACCATGAGAGACTGCAACCGAGCGCCAGGCCCACTTCTGGGCAAGATCTGGCCC 1305
Db 420 GlyHisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpPro 439
QY 1306 AGCCACAAGGGCGCGCGCGCGCAAC----- 1329
Db 440 SerTyrLysGlyArgProGlyAsnPheLeuGlnSerArgProGluProThrAlaProPro 459
QY 1330 TTCCTGCAGAGCGCGCGCGCGAGCCCGCGCGCGCGCGCGAGAGCTTCCGC-----TTC 1383
Db 460 PheLeuGlnSerArgProGluProThrAlaProGluGlnSerPheArgSerGlyVal 479
QY 1384 GAGGAGACCACCCCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1437
Db 480 GluThrThrThrProGlnLysGlnGluProIleAspLysGluLeuTyrProLeuThr 499
QY 1438 AGCTGAAGAGCCTGTTGGCAACGAGCCCTGAGCCAG 1476
Db 500 SerLeuArgSerLeuPheGlyAsnAspProSerSerGln 512

RESULT 11
US-08-463-028-8
Sequence 8, Application US/08463028
Patent No. 6610476
GENERAL INFORMATION:
APPLICANT: CHANG, Nancy T.
APPLICANT: GALLO, Robert C.
APPLICANT: WONG-STAL, Flossie
TITLE OF INVENTION: CLONING AND EXPRESSION OF HTLV-III DNA
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,028
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/693,866
FILING DATE: 23-JAN-1985
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/659,339
FILING DATE: 10-OCT-1984
ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4193US3
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 512 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: HTLV-III
FEATURE:
NAME/KEY: Protein
LOCATION: 1..512
OTHER INFORMATION: /note= "gag protein of HTLV-III"
US-08-463-028-8
Alignment Scores:
Pred. No.: 4,99e-129 length: 512
Score: 2235.00 Matches: 424
Percent Similarity: 90.64% Conservative: 41
Best Local Similarity: 82.65% Mismatches: 26
Query Match: 80.63% Indels: 22
DB: 4 Gaps: 7
US-09-475-704A-3 (1-1479) x US-08-463-028-8 (1-512)
QY 1 ATGGGCGCGCGCGCGAGCATCTGCGCGCGCGCAAGCTGGAACGCTGGAGCGCATCCGC 60
Db 1 MetGlyAlaArgAlaSerValLeuSerGlyGlyGluLeuAspArgTrpGluLysIleArg 20
QY 61 CTGCGCGCGCGCGCGCAAGAGTGTACATGATGAAGCACCTGTGTGGCGCGCGAG 120
Db 21 LeuArgProGlyGlyLysLysLysTyrLysLeuLysHisIleValTrpAlaSerArgGlu 40
QY 121 CTGAGAGATTGCGCCCTGAACCGCGCGCTGTGAGAGACGAGCGAGGGCTGCAAGCATC 180
Db 41 LeuGluArgPheAlaValAlaAsnProGlyLeuLeuGluThrSerGluGlyCysArgGlnIle 60
QY 181 ATCCGCGAGCTGACACCCCGCGCTGCAGACCGCGAGCGAGAGCTGAAGAGCTGTCAAC 240
Db 61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGluGluLeuArgSerLeuTyrAsn 80
QY 241 ACCGTGGCACCTGTACTGCTGTCACGAGAGATCGAGGTCCGCGACACCAAGAGGCC 300
Db 81 ThrValAlaThrLeuTyrCysValHisGlnArgIleGluIleLysAspThrLysGluAla 100
QY 301 CTGACAAGATGAGAGAGAGAGAGACAAGTGCAGAGAGATCCAGACGCGAGGCC 360
Db 101 LeuAspLysIleGluGluGluGlnAsnLysSerLysLysValAlaGlnGlnAla--Ala 119
QY 361 GCGGACAAGGGC-----AAGTGAGCCAGAACTACCCCATCGTGACAACTGCAG 411
Db 120 AlaAspThrGlyHisSerSerGlnValSerGlnAsnTyrProIleValGlnAsnIleGln 139
QY 412 GCGCAGATGTCACAGGCGCATCAGCCCGCGACCCCTGAACGCTGTGAGAGGTGATC 471
Db 140 GlyGlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaIleTrpValLysValVal 159
QY 472 GAGGAGAAGGCTTCAGCCCGAGGTGATCCCATGTTACCGCCCTGAGCGAGGGCGCC 531
Db 160 GluGluLysAlaPheSerProGluValIleProMetPheSerAlaLeuSerGluGlyAla 179
QY 532 ACCCCCAAGACTGAACACGATGTTGAACCGTGGCGGCGCACGAGCGCGCATGCAG 591
Db 180 ThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaIleMetGln 199
QY 592 ATGCTGAAGACCATCAACAGAGAGAGCGCGGAGTGGGACCGGTGACCCCGTGAC 651
Db 200 MetLeuLysGluThrIleAsnGluGluAlaIleGluTrpAspArgValHisProValHis 219
QY 652 GCGGCGCCCATGCGCGCGCGCGAGATGCGGAGCCCGCGGAGCGACATGCGCGGACCC 711


```

Db      220 AlaGlyProIleAlaProGlnMetSerGluProArgGlySerAspIleAlaGlyThr 239
QY      712 ACCAGCACCTGCAGAGAGATGCGCTGGATGACCAGACACCCCCATCCCGTGCGC 771
Db      240 ThrSerThrLeuGlnGlnGlnIleGlyTrpMetThrAsnProProIleProValGly 259
QY      772 GACATCTACAAGCGGTGATCATCCTGGGCTGAACAAGATCGTGGATGTACAGCCCC 831
Db      260 GluIleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerPro 279
QY      832 GTGAGCATCCTGCACATCAAGCAGGCGCCCAAGAGCCCTTCCGCGACTACGTGACCGC 891
Db      280 ThrSerIleLeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArg 299
QY      892 TTCTTCAGACCCCTGCGCGCCGAGCAGACGACCCAGAGGTGAAGAACTGGATGACCGAC 951
Db      300 PheTyrLysThrLeuArgAlaGlnGlnAlaSerGlnGlnValLysAsnTrpMetThrGlu 319
QY      952 ACCCTGCTGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTGCGCCC 1011
Db      320 ThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyPro 339
QY      1012 GCGCGCACGCTGGAGAGATGATGACCCGCTGCGCAGGCGGTGGCGCGCCCAAGCACAAG 1071
Db      340 AlaAlaThrLeuGlnGlnMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLys 359
QY      1072 GCGCGCGCTGTGCGCGCGCATGAGCCAG--GCCAACACC--AGCGTATGATGACG 1125
Db      360 AlaArgValLeuAlaGlnAlaMetSerGlnValThrAsnThrAlaThrIleMetMetGln 379
QY      1126 AAGAGCAACTTCAAGGGGCGCGCGCATGCTCAAGTGTCTCAACTGCGGCAAGAGGCGC 1185
Db      380 ArgGlyAsnPheArgAsnGlnArgLysMetValLysCysPheAsnGlyLysGlnGly 399
QY      1186 CACATCGCGCGCAACTGCGCGCGCGCGCGCAAGAGGCTGCTGGAAGTGGCGCAAGGAG 1245
Db      400 HisThrAlaArgAsnCysArgAlaProArgLysGlyCysTrpLysCysGlyLysGlu 419
QY      1246 GCGCACGATGAAGAGCTGCACCGAGCGCGCGCAACTTCTGGGCAAGATCTGGCGC 1305
Db      420 GlyHisGlnMetLysAspCysThrGlnArgGlnAlaAsnPheLeuGlyLysIleTrpPro 439
QY      1306 AGCCACAAGGGCGCGCGCGCAAC----- 1329
Db      440 SerTyrLysGlyArgProGlyAsnPheLeuGlnSerArgProGluProThrAlaProPro 459
QY      1330 TTCTGCAGAGCGCGCGCGAGCGCCACCGCGCGCGCGCGAGAGCTTCCGC-----TTC 1383
Db      460 PheLeuGlnSerArgProGluProThrAlaProProGluGlnSerPheArgSerGlyVal 479
QY      1384 GAGAGACCAACCCCGCGCGCAGAGCAGAGAGAGCAAGCAAGCAGCGCAG-----ACCTGACC 1437
Db      480 GluThrThrThrProGlnLysGlnGlnProIleAspLysGluLeuTyrProLeuThr 499
QY      1438 AGCTGAAGAGCCTGTTGCGCAACGACCCCTGAGCCAG 1476
Db      500 SerLeuArgSerLeuPheGlyAsnAspProSerSerGln 512

```

RESULT 12
US-09-952-060-33

```

; Sequence 33, Application US/09952060
; Patent No. 6733993
; GENERAL INFORMATION:
; APPLICANT: Emini, Emilio A.
; APPLICANT: Youil, Rima
; APPLICANT: Bett, Andrew J.
; APPLICANT: Chen, Ling
; APPLICANT: Kaslow, David C.
; APPLICANT: Shiver, John W.
; APPLICANT: Toner, Timothy J.
; APPLICANT: Casimiro, Danilo R.
; TITLE OF INVENTION: ENHANCED FIRST GENERATION ADENOVIRUS

```

```

; TITLE OF INVENTION: VACCINES EXPRESSING CODON OPTIMIZED HIV1-GAG, POL, NEF AND
; TITLE OF INVENTION: MODIFICATIONS
; FILE REFERENCE: 20747Y
; CURRENT APPLICATION NUMBER: US/09/952,060
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: PCT/US01/28861
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/317,814
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/279,056
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/233,180
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: tPA-gag fusion open reading frame
US-09-952-060-33

```

```

Alignment Scores:
Pred. No.: 3.79e-121 Length: 493
Score: 2106.00 Matches: 394
Percent Similarity: 92.29% Conservative: 37
Best Local Similarity: 84.37% Mismatches: 28
Query Match: 75.97% Indels: 8
Gaps: 4

```

US-09-475-704A-3 (1-1479) x US-09-952-060-33 (1-493)

```

QY      100 CTGCTGTGGCGCACGCGCGAGCTGAGAGAGTTGCCCTGAACCCCGCGCTGTGAGACC 159
Db      27 IleValTrpAlaSerArgLysLeuGlnArgPheAlaValAsnProGlyLeuLeuGluThr 46
QY      160 AGCAGGCGTGCAGACGATCATCCGCGCTGCACCCCGCGCTGCAGACCGCAGCAG 219
Db      47 SerGlnGlyCysArgGlnIleLeuGlnGlnLeuGlnProSerLeuGlnThrGlySerGlu 66
QY      220 GAGCTGAAGAGCCTGTTCAACACCGGTGCGCACCCCTGTACTGCGTGCACGAGAATCGAG 279
Db      67 GluLeuArgSerLeuTyrAsnThrValAlaThrLeuTyrCysValHisGlnLysIleAsp 86
QY      280 GTCCGCGACACCAAGAGGCGCTGGAACAAGATCGAGAGAGAGCAAGATGCGCAGCAG 339
Db      87 ValLysAspThrLysGluAlaLeuGlnLysIleGlnGlnGlnAsnLysSerLysLys 106
QY      340 AAGATCCAGCAGGCGCGAGCC-----GCCGACAAGGCAAGGTGAGCCAGAACTACCCC 393
Db      107 LysAlaGlnGlnAlaAlaAlaGlyThrGlyAsnSerSerGlnValSerGlnAsnTyrPro 126
QY      394 ATCTGCAGAACTGCAGGCGCGATGTGCACCAAGCCATCAGCCCCCGCAGCCTGAAC 453
Db      127 IleValGlnAsnLeuGlnGlnMetValHisGlnAlaIleSerProArgThrLeuAsn 146
QY      454 GCCTGGGTGAAGGTGATCGAGGAGAAAGCCTTCAGCCCGCGAGGTGATCCCATGTTCAAC 513
Db      147 AlaTrpValLysValValGlnGlnLysAlaPheSerProGluValIleProMetPheSer 166
QY      514 GCGCTGAGCAGGCGCGCACCCCGCGAGACCTGAACAGATGTTGAACACCGTGGCGGC 573
Db      167 AlaLeuSerGlnGlyAlaThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGly 186
QY      574 CACCAAGCGCGCATGCAGATGCTGAAGACACCATCAACGAGAGGCGCGCGAGTGGAGC 633
Db      187 HisGlnAlaAlaMetGlnMetLeuLysGluThrIleAsnGlnGlnAlaAlaGluTrpAsp 206
QY      634 CGCGTGACACCCGCTGCACGCGCGCCCATCGCCCCCGCGCAGATGCGCGAGCCCGCGGC 693
Db      207 ArgLeuHisProValHisAlaGlyProIleAlaProGlyGlnMetArgGluProArgGly 226

```

```

QY      694 AGCGACATCGCCGGGCAACCACAGCACCTTGCAGAGACAGATCGCCTTGATGACCAACAAC   753
        |||||
Db       227 SerAspIleAlaGlyThrThrserThrLeuGlnGlnIleGlyTrpMetThrAsnAsn   246
QY      754 CCCCCCATCCCGGTGGGGCGACATCTACAAGCGGTGATCATCTGGGGCTGAACAAGATC   813
        |||||
Db       247 ProProIleProValGlyGluIleTyrrLysArgTriPilelleuGlyLeuAsnLysIle   266
QY      814 GTGGGATGTACAGCCCCCGTAGACATCCTGTGACATCAAGACGGGCCCAAGAGCCCTTC   873
        |||||
Db       267 ValArgMetTyriserProThrserIleLeuAspIleArgGlnGlyProLysGluProphe   286
QY      874 CGGCACTACGTGGAACCGCTTCTTACAAGACCTTGCGCGCCGAGAGACACCAAGAGAGGTG   933
        |||||
Db       287 ArgAspTyrrValAspArgPheTyrrLysThrLeuArgAlaGlnAlaSerGlnGluVal   306
QY      934 AAGAACTGGATGACCGACACCCCTGCTGTGTGACAGAACGCCAACCCTGCACTGCAGAACCATC   993
        |||||
Db       307 LysAsnTrpMetThrGluThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIle   326
QY      994 CTGCGCGCTCTCGGCCCCCGCGCCAGCCTGAGAGAGATGATGACCGCTTCCAGGGCGGTG   1053
        ||:::||
Db       327 LeuLysAlaLeuGlyProAlaAlaThrLeuGlnGluMetMetThrAlaCysGlnGlyVal   346
QY      1054 GGGGGCCCCCAGCCACAAGGCCCCGTGCTGTGCGCCAGGCGATGACGCCAAGCCACAACCAAGC   1113
        |||||
Db       347 GlyGlyProGlyHisLysAlaArgValLeuAlaGluAlaMetSerGlnValThrAsnSer   366
QY      1114 -----GTGATGATGCAGAAGAGACAATTCAAGGGCCCCCGCGGCATCGTCAAGTGCTTC   1167
        ::|||
Db       367 AlaThrIleMetMetGlnArgGlyAsnPheArgAsnGlnAlaArgLysThrValLysCysPhe   386
QY      1168 AACTCGGGCAAGAGAGGGCCACATGCCCCGCAACTGCGCGCCCCCGCAAGAGAGGCTGC   1227
        |||||
Db       387 AsnCysGlyLysValGlyHisIleAlaLysAsnCysArgAlaProArgLysLysGlyCys   406
QY      1228 TGGAAAGTGGCGCAAGAGAGGGCCACAGATGAAGAACTGCACCGAGCGCCAGGCGCAACTTC   1287
        |||||
Db       407 TrpLysCysGlyLysGlnGlyHisGlnMetLysAspCysAsnGluArgGlnAlaAsnPhe   426
QY      1288 CTGGGCAAGATCTGGCCCCAGCCACACAAGGGCGCGCCCGCAACTTCTTGACAGCCGCCCC   1347
        |||||
Db       427 LeuGlyLysIleTrpProserHisLysGlyArgProGlyAsnPheLeuGlnSerArgPro   446
QY      1348 GAGCCCAACCGCCCCCGCCGCGAGAGCTTCGCGTTTC-----GAGGAGACCAACCCCGGC   1401
        |||||
Db       447 GluProThrAlaProProGluGlnSerPheArgPheGlyGlnGluLysThrThrProser   466
QY      1402 CAGAAAGCAGAGAGACCAAGAGACCGCGAG-----ACCCTGACCAAGCCTGAAGAGCCTGTTTC   1455
        |||||
Db       467 GlnLysGlnGlnProIleAspLysGlnLeuTyrrProLeuAlaSerLeuArgSerLeuPhe   486
QY      1456 GGCAACGACCCCTGAGCCAG 1476
        |||||
Db       487 GlyAsnAspProserSerGln 493

RESULT 13
US-08-127-499A-11
; Sequence 11, Application US/08127499A
; Patent No. 5510264
; GENERAL INFORMATION:
; APPLICANT: VAN ALSTYNE, Diane
; APPLICANT: SHARMA, Lawrence Rajendra
; TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED
; TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109

```

```

/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/127,499A
/ FILING DATE: 28-SEP-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BENT, Stephen A.
/ REGISTRATION NUMBER: 29,768
/ REFERENCE/DOCKET NUMBER: 51916/102/INBI
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 672-5300
/ TELEFAX: (202) 672-5399
/ TELEX: 904136
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 478 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: unknown
US-08-127-499A-11

Alignment Scores:
Pred. No.: 5,78e-117 Length: 478
Score: 2037.50 Matches: 388
Percent Similarity: 90.81% Conservative: 47
Best Local Similarity: 81.00% Mismatches: 33
Query Match: 73.50% Indels: 11
DB: 1 Gaps: 8

US-09-475-704A-3 (1-1479) x US-08-127-499A-11 (1-478)
QY 1 ATGGGCGCGCGCGCGCGCATCCTGCGCGCGCGGCAAGCTGGACGCGCTGGGAGCGCATCCGC 60
Db 1 MetGlyAlaArgAlaSerValLeuSerGlyGlyGluLeuAspArgTyrGluLysIleArg 20
QY 61 CTGGGCGCGCGCGCGGCAAGAAGTGTCTACATGATGAAGCACTGTGTGGGCCAGCCGCGAG 120
Db 21 LeuArgProGlyGlyLysLysLysTyrLysLeuLysHisIleValTyrAlaSerArgGlu 40
QY 121 CTGGAAGAAGTTCGCGCTGAACCCCGGCGCTGCTGGAGCAAGCGAGGGCTGCAAGCAGATC 180
Db 41 LeuGluArgPheAlaValAsnProGlyLeuLeuGluThrSerGluGlyCysArgGlnIle 60
QY 181 ATCGCGCCAGCTGCACACCCCGCCCTGCAGACCGGCGAGCGAGAGCTGAAGACCTGTTCAC 240
Db 61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGluGluLeuArgSerLeuTyrAsn 80
QY 241 ACCGTGGCCACCCCTGTACTGCGTGCAACGAAGATCGAGTCCCGCACCAAGAGGCC 300
Db 81 ThrValAlaThrLeuTyrCysValHisGlnArgIleGluIleLysAspThrLysGluAla 100
QY 301 CTGGAACAAGATCGAGAGGAGGAGCAGAAACAAGTGGCCAGAGAAGATCCAGAGGCCGAGGCC 360
Db 101 LeuAspLysIleGluGluGluGluGlnAsnLysSerLysLysAlaGlnGlnAla--Ala 119
QY 361 GCCGACAAGGGC-----AAGTGAAGCCAGAACTACCCCATCGTGCAAGAACTGCGAG 411
Db 120 AlaAspThrGlyHisSerSerGlnValSerGlnAsnTyrProIleValGlnAsnIleGln 139
QY 412 GGCACAGATGTGTGACACCGCCATCAGCCCCCGCACCCCTGAACGCGCTGGGTGAAGGTGATC 471
Db 140 GlyGlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTyrValLysValVal 159
QY 472 GAGGAGAAGGCTTCAGCCCCGAGGTGATCCCATGTTCACCGCGCTGAGCGAGGGCGCC 531
Db 160 GluGluLysAlaPheSerProGluValIleProMetPheSerAlaLeuSerGluGlyAla 179
QY 532 ACCCCCCAGGACCTGAACACGATGTTGAACACCGTGGCGGGCCACCAAGCGCCGATGAG 591
Db 180 ThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaIleMetGln 199

```

```

QY 592 ATGCTGAGAGACACCATCAACGAGAGCGCCGAGTGGACCGCGTGCACCCCGTGCAC 651
    |||||||:|||||||||
Db 200 MetLeuysgluThrIleasnGluGluAlaIaGluTrpAspArgValHisProValHis 219
QY 652 GCGGCGCCCATCGCCCGCGCCAGATGCGCGGAGCCCGCGCGCAGACATCGCGCGCAC 711
    |||||||:|||||||||
Db 220 AlaGlyProIleAlaProGlyGlnMetArgGluProArgGlySerAlaIaGlyThr 239
QY 712 ACCAGCACCTGCAGAGACAGATCGCGCTGATGACCAACCCCGCATCCCGTGGCC 771
    |||||||:|||||||||
Db 240 ThrSerThrLeuGlnGluGlnIleGlyTrpMetThrAsnAsnProPoleProValGly 259
QY 772 GACATCTACAAGCGGTGATCATCTGCGGCTGAACAAGATCGTGGCATGTACAGCCCC 831
    |||||||:|||||||||
Db 260 GluIleTyrIysArgTrpIleIleLeuGlyLeuAsnIleValArgMetTyrSerPro 279
QY 832 GTGAGCATCTGCACATCAAGAGAGGCGCCCAAGAGAGCCCTTCCGCGACTACGTGACCCG 891
    |||||||:|||||||||
Db 280 ThrSerIleLeuAspIleArgGlnGlyProIysGluProPheArgAspTyrValAspArg 299
QY 892 TTCCTCAAGACCTGCGCGCCGCGCAGCAGACACCCAGAGGTGAAGAACTGGATGACCGAC 951
    |||||||:|||||||||
Db 300 PheTyrIysThrLeuArgAlaGluGlnAlaSerGlnIleValIysAsnTrpMetThrGlu 319
QY 952 ACCCTGTGTGTGCAGAAAGCCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCGCC 1011
    |||||||:|||||||||
Db 320 ThrLeuLeuValGlnAsnAlaAsnProAspCysIysThrIleLeuIysAlaLeuGlyPro 339
QY 1012 GCGCGCAGCTGCAGAGATGATGACCGCTGCGCAGGCGGTGGCGCGCCAGCAAG 1071
    |||||||:|||||||||
Db 340 AlaIaThrLeuGlnGluMetMetThrAlaCysGlnGlyValGlyIlyProGlyHisIlys 359
QY 1072 GCGCGCGTGTGGCGGAGGCGATGAGCCAG--GCCAACACC--AGCGTATGATGCAG 1125
    |||||||:|||||||||
Db 360 AlaArgValLeuAlaGluAlaMetSerGlnValThrAsnThrAlaThrIleMetMetGln 379
QY 1126 AAGAGCAACTCAAGGGGCCCGCGCGCATCGTCAAGTGTCTCAACTGCGCGCAAGAGGCG 1185
    |||||||:|||||||||
Db 380 ArgGlyAsnPheArgAsnGlnArgIysMetValIysCysPheAsnIysGlyIysGlnGly 399
QY 1186 CACATCGCGCGCAACTGCGCGCGCCCGCGCAAGAGGCGTGTGAAGTGCAGGAGAG 1245
    |||||||:|||||||||
Db 400 HisThrAlaArgAsnIysArgAlaProArgIysIysGlyCysTrpIysCysGlyIysGlu 419
QY 1246 GGCACACGATGAAGAGCTGCACCGAGCGCGCAACTCTCTGGCGCAAGATC--TGG 1302
    |||||||:|||||||||
Db 420 GlnHisGlnMetIysAspCysThrGlnArgGlnAlaAsnPheLeuGlyIysIleCysLeu 439
QY 1303 CCCAGCCACAAGGCGCGCGCGCAACTCTCTGCAGAGCGCGCCGAGCCAGCCCGCCCC 1362
    |||||||:|||||||||
Db 440 ProThrArgGluGlyGln--GlyIlePhePhe-----ArgAlaAspGlnSerGlnGln 456
QY 1363 CCCGCGCAGAGCTTCGCGCTTGCAGAGACC--ACCGCGCGCCAGAGCAGAGAGAGC 1416
    |||||||:|||||||||
Db 457 ProHisHisPhePheArgAlaAspGlnSerGlnGlnProHisGlnIysArgAlaSer 475

```

RESULT 14

```

US-08-482-847-11
; Sequence 11, Application US/08482847
; Patent No. 5556757
; GENERAL INFORMATION:
; APPLICANT: VAN ALSTYNE, Diane
; APPLICANT: SHARMA, Lawrence Rajendra
; TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR
; TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR
; TITLE OF INVENTION: CNS CARRIER, ANTIBODIES THERETO, AND USES THEREOF
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.

```

```

; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,847
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/127,499
; FILING DATE: 28-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 51916/104/INBI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; US-08-482-847-11

```

Alignment Scores:

```

Pred. No.: 5.78e-117 Length: 478
Score: 2037.50 Matches: 388
Percent Similarity: 90.81% Conservative: 47
Best Local Similarity: 81.00% Mismatches: 33
Query Match: 73.50% Indels: 11
DB: 1 Gaps: 8

```

US-09-475-704A-3 (1-1479) x US-08-482-847-11 (1-478)

```

QY 1 ATGGGCGCGCGCGCCAGCATCTGCGCGGCGCAAGCTGACCGCTGGAGCGCATCCGC 60
    |||||||:|||||||||
Db 1 MetGlyAlaArgAlaSerValLeuSerGlyGlyGluLeuAspArgTrpGluIysIleArg 20
QY 61 CTGCGCGCGCGCGCAAGAGTGTACATGATGAAGCAACCTGGTGTGGCCAGCGCGAG 120
    |||||||:|||||||||
Db 21 LeuArgProGlyGlyIysIysIysIysIysIysIysIysIysIysIysIysIysIysIys 40
QY 121 CTGAGAGAAGTTCGCGCTGAACCGCGCGCTGTGAGAGACCAAGGAGGCTGCAGACATC 180
    |||||||:|||||||||
Db 41 LeuGluArgPheAlaValAsnProGlyLeuLeuGlnIleThrSerGlnIleCysArgGlnIle 60
QY 181 ATCCGCGCAGCTGCACCCCGCGCTGCAGAGCGCGGAGAGAGCTGAAGAGCTGTCAAC 240
    |||||||:|||||||||
Db 61 LeuGlyGlnLeuGlnProSerLeuGlnIleThrGlySerGlnGluLeuArgSerLeuTyrAsn 80
QY 241 ACCGTGGCCACCTGTACTGCGTGCAGAGAGATCGAGGTCCGCGCAGACCAAGAGGCC 300
    |||||||:|||||||||
Db 81 ThrValAlaThrLeuTyrCysValHisGlnArgIleGluIleIysAspThrIysGlnAla 100
QY 301 CTGAGACAAGATGAGAGAGAGAGCAACAAGTGCAGAGAGATCCAGAGCGCGAGGCC 360
    |||||||:|||||||||
Db 101 LeuAspIysIleGluGlnGlnIleAsnIysSerIysIysIysIysIysIysIysIysIys 119
QY 361 GCGACAAAGGC-----AAGGTGAGCCAGAACTACCCCATCGTGACAACTGCAG 411
    |||||||:|||||||||
Db 120 AlaAspThrGlyHisSerSerGlnValSerGlnAsnTyrProIleValGlnAsnIleGln 139
QY 412 GCGCAGATGTGCACAGGCGCATCAGCCCGCGCACCCCTGAACGCTGGGTGAAGTGTATC 471
    |||||||:|||||||||
Db 140 GlyGlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaIleValIysValIle 159
QY 472 GAGGAGAAAGGCTTCAGCCCGCAGAGGTGATCCCATGTTACCGCCCTGAGCGAGGCGCC 531

```

```

Db      |||||||
160 GluGluValAlaPheSerProGluValIleProMetPheSerAlaLeuSerGluGlyAla 179
QY      |||||||
532 ACCCCCCAGGACCTGAACAGATGTGAACACCGTGGCGCGCCAGCAGCCGCGCATGCGAG 591
Db      |||||||
180 ThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlnIleGlnAlaAlaMetGln 199
QY      |||||||
592 ATGCTGAAGGACACCATCAACGAGGAGCGCGCGAGTGGAGCCGCGTGCACCCCGTGAC 651
Db      |||||||
200 MetLeuLysGluThrIleAsnGluGlnAlaAlaGluTrpAspArgValHisProValHis 219
QY      |||||||
652 GCCGGCCCCATCGCCCCGGCCAGATGCGCGAGCCCGCGCGAGCATCGCCGCGAC 711
Db      |||||||
220 AlaGlyProIleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThr 239
QY      |||||||
712 ACCAGCACCTGACGAGCAGATCGCCTGATGACCAACCCCGCATCCCGCTGGGC 771
Db      |||||||
240 ThrSerThrLeuGlnGlnIleGlyTrpMetThrAsnAsnProIleProValGly 259
QY      |||||||
772 GACATCTACAAGCGGTGATCATCTGGGCGCTGAACAAGATCGTGGGATGTACAGCCCC 831
Db      |||||||
260 GluIleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerPro 279
QY      |||||||
832 GTGAGCATCTTGACATCAAGCAGGCGCCCAAGAGCCCTTCGCGCATACGTGAGCCGC 891
Db      |||||||
280 ThrSerIleLeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArg 299
QY      |||||||
892 TTCCTTCAAGACCTCGCGCGCGCAGCAGACACCCAGAGGTGAAGAACTGGATGACCGAC 951
Db      |||||||
300 PheTyrLysThrLeuArgAlaGluGlnAlaSerGlnGlnValLysAsnTrpMetThrGlu 319
QY      |||||||
952 ACCCTGCTGTGACAGAACGCCAACCCCGACTGCAAGACCATCTCGCGCTTCGCGCCCC 1011
Db      |||||||
320 ThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyPro 339
QY      |||||||
1012 GGCGCCAGCCTGAGAGATGATGACCGCCTGCGAGGGCGTGGCGCGCCAGCACAAG 1071
Db      |||||||
340 AlaAlaThrLeuGlnGlnMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLys 359
QY      |||||||
1072 GCCCGCGTGTGCGGAGGCGATGAGCCAG--GCCAACACC--AGCGTATGATGACAG 1125
Db      |||||||
360 AlaArgValLeuAlaGlnAlaMetSerGlnValThrAsnThrAlaThrIleMetMetGln 379
QY      |||||||
1126 AAGAGCAACTTCAAGGGCCCCCGCGCATCGTCAAGTGTTCACCTGCGGCAAGAGGGC 1185
Db      |||||||
380 ArgGlyAsnPheArgAsnGlnArgLysMetValLysCysPheAsnGlyLysGluGly 399
QY      |||||||
1186 CACATCGCCCGCAACTGCGCGCGCGCCGCAAGAGGGCTGCTGAAGTGGCGCAAGAG 1245
Db      |||||||
400 HisThrAlaArgAsnCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGlu 419
QY      |||||||
1246 GGCCACCATGATGAAGACTGCACCGAGCGCCAGGCCAATTCTCTGGGCAAGATC---TGG 1302
Db      |||||||
420 GlyHisGlnMetLysAspCysThrGlnArgGlnAlaAsnPheLeuGlyLysIleCysLeu 439
QY      |||||||
1303 CCGAGCCACAAGGGCGCGCGCGCAACTTCTGCAAGCGCGCCGAGCCAGCCCGCCCC 1362
Db      |||||||
440 ProThrArgGluGln---GlyIlePhePhe-----ArgAlaAspGlnSerGlnGln 456
QY      |||||||
1363 CCGCGCCGAGAGCTTCGCTTGCAGAGACC--ACCCCGCGCCAGAAAGCAGAGAGC 1416
Db      |||||||
457 ProHisHisPhePheArgAlaAspGlnSerGlnGlnProHisGlnLysArgAlaSer 475

```

RESULT 15

```

US-08-392-794A-2
; Sequence 2, Application US/08392794A
; Patent No. 6025141
; GENERAL INFORMATION:
; APPLICANT: HU, Yu-Wen
; TITLE OF INVENTION: IMMUNOFLUORESCENCE ASSAY FOR THE DETECTION OF
; TITLE OF INVENTION: ANTIBODIES USING RECOMBINANT ANTIGENS IN INSOLUBLE FORM
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
;

```

```

; ADDRESSEE: BROWDY AND NEWMARK
; STREET: 419 Seventh Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/392,794A
; FILING DATE: 09-DEC-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/164,789
; FILING DATE: 10-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: NEWMARK, SHERIDAN
; REGISTRATION NUMBER: 20,520
; REFERENCE/DOCKET NUMBER: HU=4A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-392-794A-2

```

Alignment Scores:

```

Pred. No.: 4 38e-116 Length: 437
Score: 2023.00 Matches: 376
Percent Similarity: 94.06% Conservative: 36
Best Local Similarity: 85.84% Mismatches: 20
Query Match: 72.98% Indels: 6
DB: 3 Gaps: 3

```

US-09-475-704A-3 (1-1479) x US-08-392-794A-2 (1-437)

```

QY      1 ATGGGCGCGCGCGCCAGCATCTGCGCGCGCGCAAGCTGAGCGCCTGGAGCGCATCCGC 60
Db      1 MetGlyAlaArgAlaSerValLeuSerGlyGlyGluLeuAspArgTrpGlyLysIleArg 20
QY      61 CTCGCGCGCGCGCGCAAGAGTGTACATGATGAAGCACCTGCTGTGGCGCAGCCGCGAG 120
Db      21 LeuArgProGlyGlyLysLysLysTyrLysLeuLysHisIleValIleTrpAlaSerArgLys 40
QY      121 CTGAGAGAAGTTGCGCCTGAACCCCGCGCTGCTGAGAGCAGCAGAGGGCTGCAAGCAGATC 180
Db      41 LeuGluArgPheAlaValAlaAsnProGlyLeuLeuGluThrSerGluGlyCysArgGlnIle 60
QY      181 ATCCGCGAGCTGCACCCCGCGCTGACAGCCGCGAGAGCTGAAGAGCTGTTCAC 240
Db      61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGlnGluLeuArgSerLeuTyrAsn 80
QY      241 ACCGTGGCACCTGTACTGCTGTCACGAGAGATGAGTCCGCGCAGCACAAGAGAGGC 300
Db      81 ThrValAlaThrLeuTyrCysValHisGlnArgIleGluIleLysAspThrLysGluAla 100
QY      301 CTGAGACAAGATGAGAGGAGGAGCAACAAGTGCAGAGAGATCCAGAGCCGAGGCC 360
Db      101 LeuAspLysIleGluGlnGlnAlaAsnLysSerLysLysValAlaGlnGlnAla--Ala 119
QY      361 GCCGACAAGGC-----AAGGTGAGCCGAAGTACCCCATCGTGACAACTGCAG 411
Db      120 AlaAspThrGlyHisSerAsnGlnValSerGlnAsnTyrProIleValGlnAsnIleGln 139
QY      412 GGCCAGATGTGTGACACGAGGCATCAGCCCGCGCACCTGAACGCGCTGTGTAAGTGATC 471

```

```
Db      140 GlyGlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValIleVal 159
QY      472 GAGAGAGAGCCCTTCAGCCCGAGGTGATCCCATGTTCAACGCCCTGAGCGAGCGCC 531
Db      160 GluGluValAlaPheSerProGluValIlePrometPheSerAlaLeuSerGluGlyAla 179
QY      532 ACCCCCGAGGACCTGAACACGATGTTGAACACCGTGGCGGCCACCGAGCCCGCATGAG 591
Db      180 ThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyHisGlnAlaAlaMetGln 199
QY      592 ATGCTGAAGGACACCATCAACGAGAGAGCGCCGAGTGGGACCGCGTGCACCCCGTGAC 651
Db      200 MetLeuValSerGluThrIleAsnGluGluAlaAlaGluTrpAspArgValHisProValHis 219
QY      652 GCCGGCCCATCGCCCGCCGAGATGGCCGAGCCCGCGGACGACATCGCCGCGCAC 711
Db      220 AlaGlyProIleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThr 239
QY      712 ACCAGCACCCCTGCAGAGACGATCGCCCTGGATGACCAACACCCCCCATCCCGTGGC 771
Db      240 ThrSerThrLeuGlnGluGlnIleGlyTrpMetThrAsnAsnProIleProValGly 259
QY      772 GACATCTACAGCGGTGATCATCCTGGCGCTGAACACAGATCGTGGGATGTACAGCCCC 831
Db      260 GluIleTyrIleAspArgTrpIleIleLeuGlyLeuAsnIleValArgMetTyrSerPro 279
QY      832 GTGAGCATCCTGCACATCAAGCAGGCGCCCAAGAGCCCTTCCGCGACTACGTGAGCCG 891
Db      280 ThrSerIleLeuAspIleArgGlnGlyProIleProPheArgAspTyrValAspArg 299
QY      892 TTCCTCAAGACCCCTGCGCGCCGAGCAGACACCCAGAGGTGAAGAACTGGATGACCGAC 951
Db      300 PheTyrIleThrLeuArgAlaGluGlnAlaSerGlnGluValIleAsnTrpMetThrGlu 319
QY      952 ACCCTGTGCTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCTCTCGGCCCC 1011
Db      320 ThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyPro 339
QY      1012 GGCGCCAGCCTGAGAGATGATGACCGCCTGCGAGGCGCTGGGCGGCCCGCCAGCCACAAG 1071
Db      340 AlaAlaThrLeuGlnGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLys 359
QY      1072 GCCCGGTGCTGGCCGAGCGGATGAGCGACGCGCCACACACAGC-----GTGATGATGAG 1125
Db      360 AlaArgValLeuAlaGluAlaMetSerGlnValThrAsnSerAlaThrIleMetMetGln 379
QY      1126 AAGAGCAACTTCAAGGGCCCCCGCGCATCGTCAAGTCTTCAACTGCGGCAAGAGGGC 1185
Db      380 ArgGlyAsnPheArgAsnGlnArgLysIleValLysCysPheAsnCysGlyLysGluGly 399
QY      1186 CACATCGCCCGCAACTGCGCGCGCCCGCCGCAAGAGGCTGCTGAAGTGGCGCAAGAG 1245
Db      400 HisThrAlaArgAsnCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGlu 419
QY      1246 GGCCACCAAGATGAAGGACTGCACCGAGCGCCAGGCCCACTTCTGGCAAGATC 1299
Db      420 GlyHisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIle 437
```

Search completed: September 1, 2005, 23:19:32
Job time : 47.739 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 1, 2005, 22:09:18 ; Search time 378.412 Seconds
(without alignments)
3078.466 Million cell updates/sec

Title: US-09-475-704A-3
Perfect score: 2772
Sequence: 1 atggcgccgcgcgcacgat.....acgacccctgagccagtaa 1479

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1774312 seqs, 393823214 residues
Total number of hits satisfying chosen parameters: 3548624

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=x1p
-Q=/cgn2_1/USPTO_spool_p/US09475704/runat_01092005_160752_19548/app_query.fasta_1.3342
-DB=Published Applications AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09475704 @CGN 1 1 233 @runat 01092005_160752_19548
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2606	94.0	492	10 US-09-899-575-17	Sequence 17, Appl
2	2464	88.9	492	18 US-10-780-507-96	Sequence 96, Appl
3	2462.5	88.8	493	18 US-10-780-507-98	Sequence 98, Appl
4	2461	88.8	492	18 US-10-780-507-97	Sequence 97, Appl
5	2448	88.3	492	16 US-10-332-413-2	Sequence 2, Appl1
6	2448	88.3	492	18 US-10-494-131-36	Sequence 36, Appl
7	2435	87.8	492	9 US-09-991-258-5	Sequence 5, Appl1
8	2435	87.8	492	18 US-10-929-234-5	Sequence 5, Appl1
9	2435	87.8	631	16 US-10-332-413-8	Sequence 8, Appl1
10	2423	87.4	492	14 US-10-339-217-110	Sequence 110, App
11	2416	87.2	492	14 US-10-339-217-109	Sequence 110, App
12	2415.5	87.1	491	14 US-10-339-217-140	Sequence 140, App
13	2407.5	86.9	491	14 US-10-339-217-104	Sequence 104, App
14	2396.5	86.5	491	14 US-10-339-217-143	Sequence 143, App
15	2382	85.9	496	14 US-10-339-217-108	Sequence 108, App
16	2381.5	85.9	487	14 US-10-339-217-106	Sequence 106, App
17	2381.5	85.9	497	14 US-10-339-217-107	Sequence 107, App
18	2344	84.6	508	14 US-10-339-217-105	Sequence 105, App
19	2343	84.5	502	10 US-09-899-575-22	Sequence 22, Appl
20	2328.5	84.0	494	14 US-10-339-217-126	Sequence 126, App
21	2297	82.9	500	14 US-10-339-217-18	Sequence 18, Appl
22	2297	82.9	500	18 US-10-780-507-47	Sequence 47, Appl
23	2281	82.3	500	18 US-10-780-507-49	Sequence 49, Appl
24	2277	82.1	498	16 US-10-325-468-22	Sequence 22, Appl
25	2266	81.7	500	16 US-10-325-468-8	Sequence 8, Appl1
26	2263	81.6	500	16 US-10-325-468-34	Sequence 34, Appl
27	2262	81.6	498	16 US-10-325-468-27	Sequence 27, Appl
28	2259	81.5	500	18 US-10-780-507-48	Sequence 48, Appl
29	2258	81.5	500	14 US-10-339-217-121	Sequence 121, App
30	2256	81.4	500	14 US-10-097-534-32	Sequence 32, Appl
31	2256	81.4	500	14 US-10-059-271-85	Sequence 85, Appl
32	2256	81.4	500	14 US-10-102-622-4	Sequence 4, Appl1
33	2256	81.4	500	17 US-10-844-658-6	Sequence 6, Appl1
34	2252.5	81.3	499	15 US-10-296-734-1	Sequence 1, Appl1
35	2252	81.2	500	16 US-10-325-468-17	Sequence 17, Appl
36	2251	81.2	500	9 US-09-968-355-26	Sequence 26, Appl
37	2251	81.2	583	9 US-09-968-355-17	Sequence 17, Appl
38	2249	81.1	500	14 US-10-224-999A-3483	Sequence 42, Appl
39	2249	81.1	500	14 US-10-224-999A-3483	Sequence 3483, Ap
40	2246.5	81.0	501	14 US-10-339-217-136	Sequence 136, App
41	2245	81.0	1350	10 US-09-952-060-35	Sequence 35, Appl
42	2245	81.0	1350	16 US-10-380-641-35	Sequence 35, Appl
43	2245	81.0	1350	17 US-10-636-730-35	Sequence 35, Appl
44	2237	80.7	557	14 US-10-339-217-144	Sequence 144, App
45	2235	80.6	512	17 US-10-634-165-8	Sequence 8, Appl1

ALIGNMENTS

RESULT 1
US-09-899-575-17
; Sequence 17, Application US/09899575
; Publication No. US20030223961A1
; GENERAL INFORMATION:
; APPLICANT: Zur Megede, Jan
; APPLICANT: Barnett, Susan W.
; APPLICANT: Egnelbrecht, Susan
; APPLICANT: van Rensburg, Estrelita Janse
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; FILE REFERENCE: PP01631.102
; CURRENT APPLICATION NUMBER: US/09/899,575
; PRIOR FILING DATE: 2001-07-05
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 17
; LENGTH: 492

```

; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-899-575-17

Alignment Scores:
Pred. No.:      3.19e-129      Length:      492
Score:          2606.00      Matches:      492
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    94.01%      Indels:      0
DB:             10      Gaps:      0

US-09-475-704A-3 (1-1479) x US-09-899-575-17 (1-492)

QY      1 ATGGGGCGCGCGCCAGATCCTGCGCGCGGCAAGCTGGAGCGCTGGAGCGCATCCGC 60
Db      1 MetGlyAlaArgAlaSerIleLeuArgGlyGlyLysLeuAspAlaTrpGluArgIleArg 20
QY      61 CTGCGCGCGCGCGGCAAGAAGTGTCTACTGATGAGCACTGGTGTGGGCCAGCGCGAG 120
Db      21 LeuArgProGlyGlyLysLysCysTyrMetMetLysHisLeuValTrpAlaSerArgGlu 40
QY      121 CTGAGAAAGTTGCGCCCTGAAACCCCGCTGTGAGACCAAGCGGCTGCAAGCAGATC 180
Db      41 LeuGluLysPheAlaLeuAsnProGlyLeuLeuGluThrSerGluGlyCysLysGlnIle 60
QY      181 ATCCGCGCAGCTGCAACCCCGCCCTGCAGACCGGCAAGGAGCTGAAGAGCTGTTCAC 240
Db      61 IleArgGlnLeuHisProAlaLeuGlnThrGlySerGluGluLeuLysSerLeuPheAsn 80
QY      241 ACCGTGGCCACCTGTACTGCGTGTGCACGAGAAGATCGAGGTCCGCGACCAAGAGAGCC 300
Db      81 ThrValAlaThrLeuTyrCysValHisGluLysIleGluValArgAspThrLysGluAla 100
QY      301 CTGAGACAAGATCGAGGAGGAGCAGACAAGTGCACAGAGAAGATCCAGCAGGCGGAGGCC 360
Db      101 LeuAspLysIleGluGluGlnGlnAsnLysCysGlnGlnLysIleGlnGlnAlaGluAla 120
QY      361 GCGGACAGGGGCAAGGTGAGCCAGAACTACCCCATCTGCAGAACTGCAGGCGCCAGATG 420
Db      121 AlaAspLysGlyLysValSerGlnAsnTyrProIleValGlnAsnLeuGlnGlyMet 140
QY      421 GTGCACCAAGGCGCATCAGCCCCCGCACCTGAACGCTGGGTGAAGGTGATCGAGGAAG 480
Db      141 ValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValIleGluGluLys 160
QY      481 GCCTTCAGCCCCGAGGTGATCCCCCATGTTCACCGCCCTGAGCGAGGCGCCACCCCCAG 540
Db      161 AlaPheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGln 180
QY      541 GACCTGAACACGATGTTGAACACCGTGGGCGGCGCACAGGCGCCCATGACAGATGCTGAAG 600
Db      181 AspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLys 200
QY      601 GACACCATCAACGAGGAGGCGCGCGAGTGGGACCGCGTGCACCCCGTGCACGCGGCGCCC 660
Db      201 AspThrIleAsnGluGlnAlaAlaGluTrpAspArgValHisProValHisAlaGlyPro 220
QY      661 ATGCGCCCCGCGCAGATGCGCGAGCGCGCGCGGAGCGAGCATGCGCGGCAACCAACGACC 720
Db      221 IleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerThr 240
QY      721 CTGCAAGAGCAGATCGCCTGATGACGACCAACCCCATCCCGGTGGGCGGACATCTAC 780
Db      241 LeuGlnGluGlnIleAlaTrpMetThrSerAsnProProIleProValGlyAspIleTyr 260
QY      781 AAGCGGTGATCATCTGGGCGCTGAACAAGATCGTGGGGATGTACAGCCCCGTGAGCATC 840
Db      261 LysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProValSerIle 280
QY      841 CTGACATCAAGCAGGCGCCCAAGAGAGCCCTTCGCGCACTACGTGACCGCTTCTTCAAG 900
Db      281 LeuAspIleLysGlnGlyProLysGluProPheArgAspTyrValAspArgPhePheLys 300

```

```

QY      901 ACCCTGCGCGCGCAGCAGACCAAGAGGTGAAGACTGGATGACCGACACCTGCTG 960
Db      301 ThrLeuArgAlaGluGlnSerThrGlnGluValLysAsnTrpMetThrAspThrLeuLeu 320
QY      961 GTGCAAGACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTGCGCCCGCGCCAGC 1020
Db      321 ValGlnAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeuGlyProGlyAlaSer 340
QY      1021 CTGAGAGAGATGATGACCGCTGCAAGGCGCGCTGAGGCGCGCCAGCCACAGGCGCGGTG 1080
Db      341 LeuGluGluMetMetThrAlaCysGlnGlyValGlyGlyProSerHisLysValArgVal 360
QY      1081 CTGCGCCGAGCGATGAGCCAGGCGCAACACCAAGCGTGTGATGATGACGAAGCACTTCAAG 1140
Db      361 LeuAlaGluAlaMetSerGlnAlaAsnThrSerValMetMetClnLysSerAsnPheLys 380
QY      1141 GCGCGCGCGCGCATCGTCAAGTGTCTTCAACTGCGGCAAGGAGGCGCACATCGCGCGCAAC 1200
Db      381 GlyProArgArgIleValLysCysPheAsnCysGlyLysGluGlnHisIleAlaArgAsn 400
QY      1201 TGCGCGCGCGCGCGCAAGAGGCGCTGTGAAGTGCAGCAAGGAGGCGCCACAGATGAAG 1260
Db      401 CysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGluGlnHisGlnMetLys 420
QY      1261 GACTGCACCGAGCGCCAGCGCAACTTCTGCGCAAGATCTGGCCAGCCACCAAGAGGCGCGC 1320
Db      421 AspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSerHisLysGlyArg 440
QY      1321 CCGCGCAACTTCTGCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 1380
Db      441 ProGlyAsnPheLeuGlnSerArgProGluProThrAlaProProAlaGluSerPheArg 460
QY      1381 TTCGAGAGACCAACCCCGCGCAAGAGAGAGCAAGAGACCGCGAGACCTGACCCAGC 1440
Db      461 PheGluGluThrThrProGlyGlnLysGlnGlnSerLysAspArgGluThrLeuThrSer 480
QY      1441 CTGAAGAGCCTGTTGCGCAACGACGACCCCTGAGCCAG 1476
Db      481 LeuLysSerLeuPheGlyAsnAspProLeuSerGln 492

RESULT 2
US-10-780-507-96
; Sequence 96, Application US/10780507
; Publication No. US20050137387A1
; GENERAL INFORMATION:
; APPLICANT: MULLINS, James I.
; APPLICANT: RODRIGO, Allen G.
; APPLICANT: LEARN, Gerald H.
; APPLICANT: LI, Fusheng
; APPLICANT: NICKLE, David C.
; APPLICANT: JENSEN, Mark A.
; TITLE OF INVENTION: ANCESTRAL AND COT VIRAL SEQUENCES, PROTEINS AND IMMUNOGENIC COMPO
; FILE REFERENCE: 16336-001320US
; CURRENT APPLICATION NUMBER: US/10/780,507
; PRIOR FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US 10/204,204
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US 60/447,586
; PRIOR FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 96
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Most recent common ancestor reconstruction of clade C gag protein
; OTHER INFORMATION: sequence

```

US-10-780-507-96

Alignment Scores:

Pred. No.:	9.5e-122	Length:	492
Score:	2464.00	Matches:	464
Percent Similarity:	97.57%	Conservative:	17
Best Local Similarity:	94.12%	Mismatches:	10
Query Match:	88.89%	Indels:	2
DB:	18	Gaps:	2

US-09-475-704A-3 (1-1479) x US-10-780-507-96 (1-492)

```
QY 1 ATGGGGCCCGCCGACATCTGCGCGCGGCAAGCTGGAGCGCTGGAGCGCATCCGC 60
Db 1 MetGlyAlaArgAlaSerIleLeuArgGlyGlyLysLeuAspThrTrpGluLysIleArg 20
QY 61 CTGCGCCCCGGCGGCAAGAGTGTCTACATGATGAAGCACTGTGTGGCCAGCCGCGAG 120
Db 21 LeuArgProGlyGlyLysLysHisTyrMetIleLysHisLeuValTrpAlaSerArgGlu 40
QY 121 CTGAGAGAGTTGCGCCCTGAACCCCGGCTGTGAGACGAGCGAGGCTGCAAGCAGATC 180
Db 41 LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrSerGluGlyCysLysGlnIle 60
QY 181 ATCCCGCAGCTGCACCCCGCCCTGTGACAGCCGCGAGCGAGAGCTGAAGCCTGTTCAAC 240
Db 61 IleLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluGluLeuLysSerLeuTyrAsn 80
QY 241 ACCGTGGCCCACTGTACTGCGTGCAGAGAGAGATGAGTCCGCGACCAAGAGAGCC 300
Db 81 ThrValAlaThrLeuTyrCysValHisGlnArgIleGluValArgAspThrLysGlnAla 100
QY 301 CTGACAAAGATCGAGAGAGAGCAAGAACTGCGCAGAGAAAGATCCAGAGCGCGAGGCC 360
Db 101 LeuAspLysIleGluGluGluGlnAsnLysSerGlnGlnLysThrGlnGlnAlaGlnAla 120
QY 361 GCCGACAAAGGCAAGGTGAGCCAGAACTACCCCATCTGCAGAACCTGCAGGGCCAGATG 420
Db 121 AlaAsp---GlyLysValSerGlnAsnTyrProIleValGlnAsnLeuGlnGlyMet 139
QY 421 GTGACCAAGGCCATCAGCCCCCGCACCTGAACGCTGGTGAAGGTGATCGAGAGAAAG 480
Db 140 ValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValIleGluGluLys 159
QY 481 GCCTTCAGCCCCGAGGTGATCCCATGTTCAACCGCCTGAGCGAGCGGCCACCCCCAG 540
Db 160 AlaPheSerProGluValIlePrometPheThrAlaLeuSerGluGlyAlaThrProGln 179
QY 541 GACCTGAACAGATGTTGAACACCGTGGCGGCGCACAGCGCCCATGAGATGCTGAAG 600
Db 180 AspLeuAsnThrMetLeuAsnThrValGlyGlnHisGlnAlaAlaMetGlnMetLeuLys 199
QY 601 GACACCATCAACGAGAGGCGCGCGAGTGGAGACCGCGTGCACCCCGTGCAGCGCGGCC 660
Db 200 AspThrIleAsnGluGluAlaAlaGluTrpAspArgLeuHisProValHisAlaGlyPro 219
QY 661 ATCGCCCCCGGCGCAGATGCGGAGCGCCCGCGCAGCAGATCGCGCGCACCAAGCAGACC 720
Db 220 ValAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerThr 239
QY 721 CTGACGAGCAGATCGCCTGATGACCAACACCCCATCCCGGTGGGGCGACATCTAC 780
Db 240 LeuGlnGluGlnIleAlaIleTrpMetThrSerAsnProIleProValGlyAspIleTyr 259
QY 781 AAGCGGTGATCATCTGGGCGCTGAACAAGATCGTGGGATGTACGCCCCGTGAGATC 840
Db 260 LysArgTyrIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProValSerIle 279
QY 841 CTGACATCAAGAGGCGCCCAAGAGAGCCCTTCGCGACTAGTGAGCCGTTCTTCAAG 900
Db 280 LeuAspIleLysGlnGlyProLysGluProPheArgAspTyrValAspArgPhePheLys 299
QY 901 ACCCTGGCGCGCGAGAGACCAAGAGAGTGAAGAACTGATGACCGACACCTGCTG 960
```

```
Db 300 ThrLeuArgAlaGluGlnAlaThrGlnAspValLysAsnTrpMetThrAspThrLeuLeu 319
QY 961 GTGACGAACGCCAACCCCGACTGTGAAGAACCATCTGCGCGCTCGGCCCGCGCCAGC 1020
Db 320 ValGlnAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeuGlyProGlyAlaThr 339
QY 1021 CTGAGAGAGATGATGACCGCTGCAGAGGCGGTGGCGGCCCGCACAGAGCCCGCTG 1080
Db 340 LeuGluGluMetMetThrAlaCysGlnGlyValGlyGlyProSerHisLysAlaArgVal 359
QY 1081 CTGCGCGAGCGCATGAGCCAGGCC---AACACCAAGCTGATGATGCAGAGCAACTTC 1137
Db 360 LeuAlaGluAlaMetSerGlnAlaAsnAsnThrAsnIleMetMetGlnArgGlyAsnPhe 379
QY 1138 AAGGCCCCCGCGCATCGTCAAGTGTCTCACTGCGGCAAGAGAGGCCACATCCCGCC 1197
Db 380 LysGlyProArgArgIleValLysCysPheAsnCysGlyLysGluGlyHisIleAlaArg 399
QY 1198 AACTGCGCGCGCGCGCGCGCAAGAGGGCTGTGAAAGTGGCGCAAGAGGGCCACAGATG 1257
Db 400 AsnCysArgAlaProAlaArgLysLysGlyCysTyrLysCysGlyLysGluGlyHisGlnMet 419
QY 1258 AAGACTGCACCGAGCGCGCCCAACTTCTGTGGCAAGATCTGGCCAGCCACAAAGGGC 1317
Db 420 LysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSerHisLysGly 439
QY 1318 CGCCCCGCACTTCTCTGACAGAGCGCGCCGAGCCACCGCCCCCGCGAGAGCTTC 1377
Db 440 ArgProGlyAsnPheLeuGlnSerArgProGluProThrAlaProProAlaGluSerPhe 459
QY 1378 CGCTTCGAGAGAGACCAACCCCGCGCAAGAGAGAGCAAGAGCGCGAGACCTGACC 1437
Db 460 ArgPheGluGluThrThrProAlaProLysGlnGluProLysAspArgGluProLeuThr 479
QY 1438 AGCCTGAAGAGCCTGTTGCGCAACGACCCCTGAGCCAG 1476
Db 480 SerLeuLysSerLeuPheGlySerAspProLeuSerGln 492
```

RESULT 3

```
US-10-780-507-98
; Sequence 98, Application US/10780507
; Publication No. US20050137387A1
; GENERAL INFORMATION:
; APPLICANT: MULLINS, James I.
; APPLICANT: RODRIGO, Allen G.
; APPLICANT: LEARN, Gerald H.
; APPLICANT: LI, Fusheng
; APPLICANT: NICKLE, David C.
; APPLICANT: JENSEN, Mark A.
; TITLE OF INVENTION: ANCESTRAL AND COT VIRAL SEQUENCES, PROTEINS AND IMMUNOGENIC COMPO
; FILE REFERENCE: 16336-001320US
; CURRENT APPLICATION NUMBER: US/10/780,507
; PRIOR FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US 10/204,204
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US 60/447,586
; PRIOR FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 98
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Minimum of means center of tree reconstruction of clade C gag pro
; OTHER INFORMATION: tein sequence
US-10-780-507-98
```

Alignment Scores:

Pred. No.:	1,14e-121	Length:	493
Score:	2462.50	Matches:	461
Percent Similarity:	97.36%	Conservative:	19
Best Local Similarity:	93.51%	Mismatches:	12
Query Match:	88.83%	Indels:	1
DB:	18	Gaps:	1

US-09-475-704A-3 (1-1479) x US-10-780-507-98 (1-493)

QY	1	ATGGGCGCGCCGCGCCAGCATCTCGCGCGCGGCAAGCTGGACCGCTGGAGCGCATCCGC	60
Db	1	MetGlyAlaArgAlaSerIleLeuArgGlyGlyLysLeuAspThrTrrpGlyLysIleArg	20
QY	61	CTGCGCGCGCGCGCGCAAGAAGTGTCTACATGATGAAGCACCTGGTGTGGGCGAGCCGGAG	120
Db	21	LeuArgProGlyGlyLysLysHisTyrMetLeuLysHisIleValTrrpAlaSerArgGlu	40
QY	121	CTGGAGAAGTTCGCCCTGTAACCCCGGCGCTGCTGGAGACCGAGGGGCTGCAAGCAGATC	180
Db	41	LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrSerGlyCysLysGlnIle	60
QY	181	ATCCGCCACGCTGCACCCCGCCCTGCAGACCGGCGAGCGAGAGCTGAAGAGCTGTTCAC	240
Db	61	MetLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluGluLeuArgSerLeuTyrAsn	80
QY	241	ACCGTGGCCACCCCTGTACTGCGGTGCACGACGAGAAGATCGAGGTCCGCGACACCAAGAGCC	300
Db	81	ThrValAlaThrLeuTyrCysValHisGlyLysIleGluValArgAspThrLysGluAla	100
QY	301	CTGCACACAGATCGAGAGGAGCAGACAAAGTCCAGCAGACAGATCCAGCAGCCGAGGCC	360
Db	101	LeuAspLysIleGlyGluGluGlnAsnLysSerGlnGlnLysThrGlnGlnAlaGluAla	120
QY	361	GCCGACAAAGGGCAAGGTAGCCAGACTACCCCATCGTGCAGAACCTGCAGGGCCAGATG	420
Db	121	AlaAlaAspGlyLysValSerGlnAsnTyrProIleValGlnAsnLeuGlnGlyGlnMet	140
QY	421	GTGCACCAAGGCCATCAGCCCCCGCACCTGTGAACGCTGGGTGAAGGTGATCGAGAGAG	480
Db	141	ValHisGlnAlaIleSerProArgThrLeuAsnAlaTrrpValLysValIleGluGluLys	160
QY	481	GCGTTCAGCCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGCGCCACCCCCAG	540
Db	161	AlaPheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGln	180
QY	541	GACCTGAACACGATGTTGAACACCGTGGGCGGCCACAGCGCCGCATGACATGCTGAAG	600
Db	181	AspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLys	200
QY	601	GACACCATCAACGAGGAGGCGCGCGAGTGGGACCGGCTGCACCCCGTGCACGCGGCCCC	660
Db	201	AspThrIleAsnGluGluAlaAlaGluTrrpAspArgLeuHisProValHisAlaGlyPro	220
QY	661	ATCGCCCCCGCCAGATGCGCGAGCGCCCGCGGCGAGCGACATCGCCGCGACCAACGACCC	720
Db	221	ValAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerThr	240
QY	721	CTGCAGGAGCAGATCGCGCTGGATGACCAACCAACCCCCCAATCCCGTGGGCGACATCTAC	780
Db	241	LeuGlnGluGlnIleAlaTrrpMetThrSerAsnProProValProValGlyAspIleTyr	260
QY	781	AAAGCGGTGATCATCTGGGCTGTAAACAAGATCGTGGGATGTACAGCCCGGTGAGCATC	840
Db	261	LysArgTrrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProValSerIle	280
QY	841	CTGCACATCAAGCAGGGCCCCAAGGAGCCCTTCGCGACTACGTGACCCGCTTCTTCAAG	900
Db	281	LeuAspIleLysGlnGlyProLysGluProPheArgAspTyrValAspArgPhePheLys	300
QY	901	ACCTGCGCGCGGAGCAGAGCACCCAGAGAGGTGAAGAACTGGATGACCGACACCCCTGCTG	960
Db	301	ThrLeuArgAlaGluGlnAlaThrGlnAspValLysAsnTrrpMetThrAspThrLeuLeu	320

[illegible]

RESULT 4

```

US-10-780-507-97
; Sequence 97, Application US/10780507
; Publication No. US20050137387A1
; GENERAL INFORMATION:
; APPLICANT: MULLINS, James I.
; APPLICANT: RODRIGO, Allen G.
; APPLICANT: LEARN, Gerald H.
; APPLICANT: LI, Fusheng
; APPLICANT: NICKLE, David C.
; APPLICANT: JENSEN, Mark A.
; TITLE OF INVENTION: ANCESTRAL AND COT VIRAL SEQUENCES, PROTEINS AND IMMUNOGENIC COMPO
; FILE REFERENCE: 16336-001320US
; CURRENT APPLICATION NUMBER: US/10/780,507
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US 10/204,204
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US 60/447,586
; PRIOR FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 97
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Least squares center of tree reconstruction of clade C gag protei
; OTHER INFORMATION: n sequence
US-10-780-507-97

```

Alignment Scores: 1.37e-121 Length: 492
Pred. No.:

Score:	2461.00	Matches:	462
Percent Similarity:	97.57%	Conservative:	19
Best Local Similarity:	93.71%	Mismatches:	10
Query Match:	88.78%	Indels:	2
DB:	18	Gaps:	2
US-09-475-704A-3 (1-1479) x US-10-780-507-97 (1-492)			
QY	1 ATGGGCGCCCGCGCCAGCATCTCTGGCGCGCGGCAAGCTGAGCGCTGGAGCCGATCCGC	60	
Db	1 MetGlyAlaArgAlaSerIleLeuArgGlyGlyLeuAspThrTrpGluysIleArg	20	
QY	61 CTGGCGCCCGCGCGCAAGAGTGTACATGATGAAGCACCTGTGTGGCCAGCCGCGAG	120	
Db	21 LeuArgProGlyGlyLysLysHisTyrMetLeuLysHisLeuValTrpAlaSerArgGlu	40	
QY	121 CTGGAGAAGTTCGCTTGAACCCCGGCTGTGTGAGACCCAGCGGCTGTCAAGATC	180	
Db	41 LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrSerGluGlyCysLysGlnIle	60	
QY	181 ATCCGCCAGCTGCAACCCCGCTGACAGCCGCGAGAGAGTGAAGAGCTGTTCAC	240	
Db	61 MetLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluGluLeuArgSerLeuTyrAsn	80	
QY	241 ACCGTGGCCACCTGTACTGCGGTGCACGAGAATCGAGGTCCGCGACACCAAGAGGCC	300	
Db	81 ThrValAlaThrLeuTyrCysValHisGluLysIleGluValAlaArgAspThrLysGluAla	100	
QY	301 CTGGACAAGATCGAGGAGGAGGACAAGCAAGTCCAGCAGAGAATCCAGAGCCGAGGCC	360	
Db	101 LeuAspLysIleGluGluGluGlnAsnLysSerGlnGlnLysThrGlnGlnAlaGluAla	120	
QY	361 GCGGACAAGGGCAAGGTAGCCAGAACTACCCCATCTGTGACAACTGACAGGCCAGATG	420	
Db	121 AlaAsp--GlyLysValSerGlnAsnTyrProIleValGlnAsnLeuGlnGlyGlnMet	139	
QY	421 GTGCACCAAGGCCATCAGCCCCCGGACCTGAAGCGCTGGTGAAGGTGATCGAGGAAG	480	
Db	140 ValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValIleGluGluLys	159	
QY	481 GCCTTCAGCCCCGAGGTGATCCCATGTTCAACCGCTGAGCGAGGGCGGCCACCCCGAG	540	
Db	160 AlaPheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGln	179	
QY	541 GACCTGAACACGATGTTGAACACCGTGGCGGCGCACCGCCCATGACAGTCTGAAG	600	
Db	180 AspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLys	199	
QY	601 GACACCATCAACGAGGAGCGCGCGAGTGGAGCCGCTGACCCCGTGCACGCGCGCCCG	660	
Db	200 AspThrIleAsnGluGluAlaIleGluTrpAspArgLeuHisProValHisAlaGlyPro	219	
QY	661 ATGCGCCCCCGCGAGATGCGCGGAGCCCCGCGGAGCGACATGCGCGCACCGACACC	720	
Db	220 ValAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerThr	239	
QY	721 CTGCAGAGCAGATCGCTGTGATGACCAAGCAACCCCATCTCCCGTGGCGCATCTAC	780	
Db	240 LeuGlnGluGlnIleAlaTrpMetThrSerAsnProProValProValGlyAspIleTyr	259	
QY	781 AAGCGGTGATCATCTCGGCTTGAACAAGATCGTGGAGTGTACAGCCCGCTGAGCATC	840	
Db	260 LysArgTrpIleIleLeuGlyLeuAsnLysIleValAlaArgMetTyrSerProValSerIle	279	
QY	841 CTGCACATCAAGAGGGGCCCCCAAGAGAGCCCTTCGCGAGTACGAGCCGCTTCTCAAG	900	
Db	280 LeuAspIleLysGlnGlyProLysGluProPheArgAspTyrValAlaAspArgPhePheLys	299	
QY	901 ACCCTGCGCGCGGAGCAGAGCACCCAGAGGTGAAGAAGTGTGATGACCGACACCTGTG	960	
Db	300 ThrLeuArgAlaGluGlnAlaThrGlnAspValLysAsnTrpMetThrAspThrLeuLeu	319	
QY	961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTTCGCGCGCGCGCCAGC	1020	

Db	320	ValGlnAsnAlaAsnProAbpCysIysThrIleLeuArgAlaLeuGlyProGlyAlaThr	339
QY	1021	CTGGAGAGATGATGACCCGCTGCCAGGGCGTGGGGCCCAAGCCACAAGGCGCGTG	1080
Db	340	LeuGIuGIuMetMetThrAlaCysGInGlyValGlyGlyProGlyHisIysAlaArgVal	359
QY	1081	CTGGCCGAGGCGATGAGCCAGGCC--AAACACCAAGCGTGAATGATGCAGAAGCAACTTC	1137
Db	360	LeuAlaGIuAlaMetSerGInAlaAsnAenThraSnIleMetMetGInArgSerAenPhe	379
QY	1138	AAGGGCCCCCGCGCATCCGTCAGTGCTTCAACTGCGGCAAGAGGGCCATCGCCCC	1197
Db	380	LysGIyProIysArgIleValIysCysPheAsnCysGIyLysGIuGIyHisIleAlaArg	399
QY	1198	AACTGCCGCGCCCCCGCACAAGGGCTGCTGGAAGTGGGGCAAGAGGGCCACCATG	1257
Db	400	AsnCysArgAlaProArgIysLysGIyCysTrpIysCysGIyLysGIuGIyHisGInMet	419
QY	1258	AAGGACTGCACCGAGCGCCGACCACTTCCTGGGCAAGATCTGGCCCAAGAGGC	1317
Db	420	LysAspCysThrGIuArgGInAlaAsnPheLeuGIyLysIleTrpProSerHisIysGIy	439
QY	1318	CGCCCCGGCAACTTCCTGCAGAGCGCGCCGACCCACCGCCCCCGCGAGAGCTTC	1377
Db	440	ArgProGIyAsnPheLeuGInSerArgProGIuProThraIaProProAlaGIuSerPhe	459
QY	1378	CGCTTCGAGAGACCAACCCCGGCGCAGAAGCAAGAGCAAGAGCCGCGACCTGACC	1437
Db	460	ArgPheGIuGIuThrThrProAlaProIysGInGIuProIysAspArgGIuProLeuThr	479
QY	1438	AGCCTGAAGACCTGTTCGGCAACGACCCCTTGAGCCAG	1476
Db	480	SerLeuIysSerLeuPheGIySerAspProLeuSerGIn	492
RESULT 5			
US-10-332-413-2			
; Sequence 2, Application US/10332413			
; Publication No. US2004011660A1			
; GENERAL INFORMATION:			
; APPLICANT: Johnstone, Robert Edward			
; APPLICANT: Swannstrom, Ronald Ivar			
; APPLICANT: Morris, Lynn			
; APPLICANT: Karim, Salim Abdool			
; APPLICANT: Williams, Carolyn			
; TITLE OF INVENTION: Process for the Selection of HIV-1 Subtype C Isolates, Selected H			
; FILE REFERENCE: 45669-281993			
; CURRENT APPLICATION NUMBER: US/10/332,413			
; CURRENT FILING DATE: 2003-09-12			
; PRIOR APPLICATION NUMBER: PCT/IB01/01208			
; PRIOR FILING DATE: 2001-07-09			
; PRIOR APPLICATION NUMBER: US 60/216,995			
; PRIOR FILING DATE: 2000-07-07			
; PRIOR APPLICATION NUMBER: ZA 2000/3437			
; PRIOR FILING DATE: 2000-07-10			
; PRIOR APPLICATION NUMBER: ZA 2000/4924			
; PRIOR FILING DATE: 2000-09-15			
; NUMBER OF SEQ ID NOS: 32			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 2			
; LENGTH: 492			
; TYPE: PRT			
; ORGANISM: Human immunodeficiency virus type 1			
US-10-332-413-2			
Alignment Scores:			
Pred. No.: 6.61e-121 Length: 492			
Score: 248.00 Matches: 461			
Percent Similarity: 97.36% Conservative: 19			
Best Local Similarity: 93.51% Mismatches: 11			
Query Match: 88.31% Indels: 2			
DB: 16 Gaps: 2			

US-09-475-704A-3 (1-1479) x US-10-332-413-2 (1-492)

```
QY 1 ATGGGCGCCGCGCCAGCATCTGCGCGCGGCAAGCTGACGCGCTGGGAGCGCATCCGC 60
Db 1 MetGlyAlaArgLaserIleLeuArgGlyLysLeuAspLysTrpGlyLysIleArg 20
QY 61 CTGCGCCCGCGCGCAAGAGTGCTACATGATGAAGCACTGTGTGGCCAGCGCGAG 120
Db 21 LeuArgProGlyLysLysHisIleTyrMetLeuLysHisIleValTrpAlaSerArgGlu 40
QY 121 CTGAGAAGTTGCGCTGAACCCCGCGCTGTGAGACCGAGCGAGCTGAAGCTTCAAC 180
Db 41 LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrSerGlyCysLysGlnIle 60
QY 181 ATCCGCGAGCTGACCCCGCGCTGACAGCCGCGAGCGAGCTGAAGCTGTTCAC 240
Db 61 MetLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluGlyLeuLysSerLeuTyrAsn 80
QY 241 ACCGTGCGCACCCCTGTACTGCGTGACAGAGAAGATCGAGTCCGCGACACCAAGAGGCC 300
Db 81 ThrValAlaThrLeuTyrCysValHisGlyLysIleGluValArgAspThrLysGluAla 100
QY 301 CTGACAAGATCGAGAGGAGCAGAACAGAGTGCCAGCAGAGAAGATCCAGCGCGAGGCC 360
Db 101 LeuAspLysIleGlyLysGlnGluGlnAsnLysCysGlnGlnLysThrGlnGlnAlaLysAla 120
QY 361 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGACAGAGCGCGCCAGATG 420
Db 121 AlaAsp---GlyLysValSerGlnAsnTyrProIleValGlnAsnLeuGlnGlyMet 139
QY 421 GTGACACGAGGCCATCAGCCCGCGACCGCTGAACGCTGGTGAAGGTGATCGAGAGAAG 480
Db 140 ValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValIleGlyLys 159
QY 481 GCCTTCAGCCCGGAGGTGATCCCATGTTCAACGCGCTGAGCGAGCGCGCACCCCGAG 540
Db 160 AlaPheSerProGluValIleProMetPheThrAlaLeuSerGlyLysAlaThrProGln 179
QY 541 GACCTGAACAGATGTTGAACACCGTGGCGCGCCACAGCGCGCCATGACAGATGCTGAAG 600
Db 180 AspLeuAsnThrMetLeuAsnThrValGlyGlnHisGlnAlaAlaMetGlnMetLeuLys 199
QY 601 GACACCATCAACGAGAGCGCGCGCGAGTGGAGCCGCGTGACCCCGTGACGCGCGCCC 660
Db 200 AspThrIleAsnGlyLysAlaIleGluTrpAspArgValHisProValHisAlaGlyPro 219
QY 661 ATCGCCCGCGCGCAGATGCGCGAGCGCGCGCGGACGACATCGCGCGACACCAAGCAC 720
Db 220 IleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerThr 239
QY 721 CTGACGAGCAGATCGCGCTGATGACAGCAACCCCGCATCCCGGTGGCGGACATGTAC 780
Db 240 LeuGlnGlnGlnIleAlaTrpMetThrSerAsnProProIleProValGlyAspIleTyr 259
QY 781 AAGCGGTGATCATCTGGGCGCTGAACAAGATCGTGGGATGTACAGCCCGCTGAGCATC 840
Db 260 LysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProValSerIle 279
QY 841 CTGACATCAAGCAGGGGCGCCAGAGAGCCCTTCGCGCACTACGTGACCGCTTCTTCAAG 900
Db 280 LeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhePheLys 299
QY 901 ACCCTGCGCGCGCGAGCAGACACCCAGAGAGGTGAAGAACTGGATGACCGACACCTGTG 960
Db 300 ThrLeuArgAlaGlnGlnAlaThrGlnGluValLysAsnTrpMetThrAspThrLeuLeu 319
QY 961 GTGCAGAACGCCAACCCCGCACTGCAAGACCATCTCGCGCGCTCTCGGCGCGCGCGCCAGC 1020
Db 320 ValGlnAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeuGlyProGlyAlaThr 339
QY 1021 CTGAGAGATGATGACCGCGCTGCGAGGCGGTGGCGGCGCCAGCCACAGGCGCGGTG 1080
```

```
Db 340 LeuGlnGlnMetMetThrAlaCysGlnGlyValGlyLysProGlyHisLysAlaArgVal 359
QY 1081 CTGGCCGAGGCGCATGAGCCAGGCCAACACCC---AGCGTATGATGCAGAGACCACTTC 1137
Db 360 LeuAlaGlnAlaMetSerGlnAlaAsnSerGlyAsnIleMetMetGlnArgSerAsnPhe 379
QY 1138 AAGGAGCCCGCGCGCATCGTCAAGTGCTTCAACTGCGGCAAGAGGGCCACATCGCCCGC 1197
Db 380 LysGlyProArgArgIleValLysCysPheAsnCysGlyLysGlnGlyHisIleAlaArg 399
QY 1198 AACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1257
Db 400 AsnCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGlnGlyHisGlnMet 419
QY 1258 AAGGACTGCACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1317
Db 420 LysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSerHisLysGly 439
QY 1318 CGCCCGCGCGCACTTCTGACAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
Db 440 ArgProGlyAsnPheLeuGlnAsnArgProGluProThrAlaProProAlaGluSerPhe 459
QY 1378 CGCTTCGAGGAGACCAACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1437
Db 460 ArgPheGlnGluThrThrProAlaProLysGlnGluProIleGluArgGluProLeuThr 479
QY 1438 AGCCTGAAGAGCCTGTTGCGCAACGACCCCGCTGAGCGCAG 1476
Db 480 SerLeuLysSerLeuPheGlySerAspProLeuSerGln 492
```

RESULT 6

```
US-10-494-131-36
; Sequence 36, Application US/10494131
; Publication No. US20050176929A1
; GENERAL INFORMATION:
; APPLICANT: Williamson, Carolyn
; APPLICANT: Van Harmelen, Joanne Heidi
; APPLICANT: Gray, Clive Maurice
; APPLICANT: Kaurin, William
; TITLE OF INVENTION: HIV-1 Subtype Isolate Regulatory/Accessory Genes, and
; TITLE OF INVENTION: Modifications and Derivatives Thereof
; FILE REFERENCE: 45669-300571
; CURRENT APPLICATION NUMBER: US/10/494,131
; PRIOR FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: PCT/IB02/04550
; PRIOR FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: ZA 2001/8978
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 36
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-494-131-36
```

Alignment Scores:

Pred. No.:	6.61e-121	Length: 492
Score:	2448.00	Matches: 461
Percent Similarity:	97.36%	Conservative: 19
Best Local Similarity:	93.51%	Mismatches: 11
Query Match:	88.31%	Indels: 2
DB:	18	Gaps: 2

US-09-475-704A-3 (1-1479) x US-10-494-131-36 (1-492)

```
QY 1 ATGGGCGCCGCGCGCCAGCATCTGCGCGCGGCAAGCTGAGCGCTGGAGCGCATCCGC 60
Db 1 MetGlyAlaArgLaserIleLeuArgGlyLysLeuAspLysTrpGlyLysIleArg 20
QY 61 CTGCGCCCGCGCGCAAGAGTGCTACATGATGAAGCACTGTGTGGCCAGCGCGAG 120
```

Db 21 LeuArgProGlyGlyLysLysHisTyrMetLeuLysHisIleValTrrPalSeraArgGlu 40
QY 121 CTGAGAAAGTTGCGCTTGAACCCCGGCTGTGAGAACCGAGGCGTGCAGACAGATC 180
Db 41 LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrSerGluGlyCysLysGlnIle 60
QY 181 ATCCCGCAGCTGCACCCCGCCCTGCAGACCGGACGAGAGCTGAAGACCTGTTCAC 240
Db 61 MetLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluLysSerLeuTyrAsn 80
QY 241 ACCGTGGCCACCTGTACTGCGTGCAAGAGATGAGGTCCGCGACACCAAGAGGCGC 300
Db 81 ThrValAlaThrLeuTyrCysValHisGluLysIleGluValArgAspThrLysGluAla 100
QY 301 CTGGACAAGATCGAGAGAGAGCAGAACTGCGCAGAGAGATCCAGAGCGCGAGGCC 360
Db 101 LeuAspLysIleGluGluGluGlnAsnLysCysGlnGlnLysThrGlnGlnAlaLysAla 120
QY 361 GCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATGCTGAGAACTGACAGGCGCAGATG 420
Db 121 AlaAsp--GlyLeValSerGlnAsnTyrProIleValGlnAsnLeuGlnGlyGlnMet 139
QY 421 GTGCACGAGGCGCATGACCCCGCAGACCTGAACGCGTGGTGAAGGTGATCGAGAGAA 480
Db 140 ValHisGlnAlaIleSerProArgThrLeuAsnAlaTrrValLysValIleGluGluLys 159
QY 481 GCCTTCAGCCCGCAGGTGATCCCATGTTCACCGCCCTGAGCGAGGCGCCACCCCGCAG 540
Db 160 AlaPheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGln 179
QY 541 GACCTGAACACGATGTTGAACACCGTGGGCGCGCCAGCAGCGCGCATGACATGCTGAAG 600
Db 180 AspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLys 199
QY 601 GACACCATCAACGAGAGAGCGCGCGCAGTGGAGCCGCGTGACCCCGTGACGCGCGCCCC 660
Db 200 AspThrIleAsnGluGluAlaIleGluTrrAspArgValHisProValHisAlaGlyPro 219
QY 661 ATGCCCCCGCGCAGATGCGCGAGCGCGCGCGCAGGACATCGCGCGCACACCGACACC 720
Db 220 IleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerThr 239
QY 721 CTGCAGAGAGAGATCGCCTGATGACCAAGCAACCCCGCATCCCGTGGCGGACATCTAC 780
Db 240 LeuGlnGluGlnIleAlaTrrMetThrSerAsnProIleProValGlyAspIleTyr 259
QY 781 AAGCGGTGATCATCTCTGGGCGCTGAACAAGATCGTCCGGATGTAACGCCCGTGACATC 840
Db 260 LysArgTrrIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProValSerIle 279
QY 841 CTGCACATCAAGCAGGCGCGCCCAAGAGCGCTTCCCGACTACGTTGACCGCTTCTTCAAG 900
Db 280 LeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhePheLys 299
QY 901 ACCCTGCGCGCGCAGACAGACACCGAGAGGTGAAGAACTGGATGACCGACCCCTGCTG 960
Db 300 ThrLeuArgAlaGluGlnAlaThrGlnGluValLysAsnTrrMetThrAspThrLeuLeu 319
QY 961 GTGCAGAACGCCAACCCCGACTGCAAGACCATCTCGCGCGCTTTCGCGCGCGCGCGCAGC 1020
Db 320 ValGlnAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeuGlyProGlyAlaThr 339
QY 1021 CTGGAGGAGATGATGACCGCCTGCGCAGGGCGGTGGGCGCGCCCGACCGACAGCGCGCGCTG 1080
Db 340 LeuGluGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLysAlaArgVal 359
QY 1081 CTGCGCGGAGGCGCATGACCGCAGCGCCCAACAC--AGCGTGAATGATGACAGAGCAACTTC 1137
Db 360 LeuAlaGluAlaMetSerGlnAlaAsnSerGlyAsnIleMetMetGlnArgSerAsnPhe 379
QY 1138 AAGGCGCGCGCGCGCATGCTCAAGTGTCTTCAACTGCGCGGCAAGGAGGCGCACATGCGCGCGC 1197
Db 380 LysGlyProArgArgIleValLysCysPheAsnCysGlyLysGluGlyHisIleAlaArg 399

QY 1198 AACTGCGCGCGCGCGCGCGCAGAGAGGCTGCTGAAGTGCAGGCAAGAGGCGCACAGATG 1257
Db 400 AsnCysArgAlaProArgLysLysGlyCysTrrPlyCysGlyLysGluGlyHisGlnMet 419
QY 1258 AAGACTGCACCGCAGCGCGCAGCGCAACTCTCTGGCGCAAGATCTGGCGCCAGCCCAAGGCG 1317
Db 420 LysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrrProSerHisLysGly 439
QY 1318 CGCCCCGGCAACTCTCTGCAGAGCGCGCGCGCCAGCCCGACCGCGCGCGCAGACTTC 1377
Db 440 ArgProGlyAsnPheLeuGlnAsnArgProGluProThrAlaProProAlaGluSerPhe 459
QY 1378 CGCTTCGAGAGACCAACCCCGCGCCAGAGCAGAGCAGAGCAGCAGCGCGAGCTGACC 1437
Db 460 ArgPheGluGluThrThrProAlaProLysGlnGluProIleGluArgGluProLeuThr 479
QY 1438 AGCCTGAAGAGCTGTTCGCGCAACGACCCCGTGAGCCAG 1476
Db 480 SerLeuLysSerLeuPheGlySerAspProLeuSerGln 492

RESULT 7
US-09-991-258-5
; Sequence 5, Application US/09991258
; Patent No. US20020141975A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Robert
; APPLICANT: Keith, Paula
; APPLICANT: Dryga, Sergey
; APPLICANT: Caley, Ian
; APPLICANT: Maughan, Maureen
; APPLICANT: Johnston, Robert
; APPLICANT: Davis, Nancy
; APPLICANT: Swannstrom, Ronald
; TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOMES WITH MODIFIED HIV GENES FOR USE
; TITLE OF INVENTION: VACCINES
; FILE REFERENCE: 01113.0001U3
; CURRENT APPLICATION NUMBER: US/09/991,258
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 09/902,537
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,995
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. US20020141975A1e =
; OTHER INFORMATION: synthetic construct
US-09-991-258-5

Alignment Scores:
Pred. No.: 3.19e-120 Length: 492
Score: 2435.00 Matches: 458
Percent Similarity: 96.96% Conservative: 20
Best Local Similarity: 92.90% Mismatches: 13
Query Match: 87.84% Indels: 2
DB: 9 Gaps: 2

US-09-475-704A-3 (1-1479) x US-09-991-258-5 (1-492)

QY 1 ATGGGCGCGCGCGCAGATCCTGCGCGGCGCAAGCTGAGCGCTGGAGCGCATCCGC 60
Db 1 MetaAlaAlaArgAlaSerIleLeuArgGlyGluLysLeuAspLysTrrGluLysIleArg 20
QY 61 CTGCGCGCGCGCGCAGAGAGTGTCTACATGATGAGACACTGTGTGGCGCAGCGCGAG 120
Db 21 LeuArgProGlyGlyLysLysHisTyrMetLeuLysHisIleValTrrPalSeraArgGlu 40
QY 121 CTGAGAAAGTTGCGCGCTGAACCCCGGCTGTGAGAACCGAGGCGTGCAGACAGATC 180

Db 41 LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrSerGluGlyCysLysGlnIle 60
QY 181 ATCCGGCAGCTGCACCCCGCCCTGCAGACCGGCAGCGAGCTGAAGACCTGTTCAAC 240
Db 61 MetLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluGluLeuLysSerLeuTyrAsn 80
QY 241 ACCGTGGCCACCTGTACTGCTGCTGCAGAGAGATCGAGGTCCGGCAGACCAAGGAGGCC 300
Db 81 ThrValAlaThrLeuTyrCysValHisGluLysIleGluValArgAspThrLysGluAla 100
QY 301 CTGGACAAGATCGAGAGAGAGAGACAAGTGCAGAGAGATCCAGAGAGAGAGAGAGGCC 360
Db 101 LeuAspLysIleGluGluGluGlnAsnLysCysGlnGlnLysThrGlnGlnAlaLysAla 120
QY 361 GCCGACAGGGCAAGGTGAGCCAGAACTACCCCATCTGCAGAACTGCAGAGGCCAGATG 420
Db 121 AlaAsp---GlyLysValSerGlnAsnTyrProIleValGlnAsnLeuGlnGlyGlnMet 139
QY 421 GTGCACACGAGCCATCAGCCCGCCAGCCCTGAACGCTGGGTGAAGGTGATCGAGAGAA 480
Db 140 ValHisGlnAlaIleSerProArgThrLeuAsnAlaThrValLysValIleGluGluLys 159
QY 481 GCCTTCAGCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGCGCCACCCCCAG 540
Db 160 AlaPheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGln 179
QY 541 GACCTGAACACGATGTTGAACACCGTGGCGGGCCACAGGCGCCCATGACAGATGCTGAAG 600
Db 180 AspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLys 199
QY 601 GACACCATCAACGAGAGAGGCGCGCGAGGTGGAGCCGCTGCACCCCGTCACGCGGCCCC 660
Db 200 AspThrIleAsnGluGluAlaIleGluThrPaspArgLeuHisProValHisAlaGlyPro 219
QY 661 ATCGCCCCCGGCAGATGCGCGAGCGCCCGCGAGACATCGCGCGCCAGACCAACCAACC 720
Db 220 IleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerThr 239
QY 721 CTGCAGAGAGAGATGCGCTGGATGACCAACACCCCATCCCGTGGCGGAGCATCTAC 780
Db 240 LeuGlnGluGlnIleAlaThrMetThrSerAsnProProIleProValGlyAspIleTyr 259
QY 781 AAGCGGTGATCATCTGGGCTTGAACAAGATCGTGGGATGTACGCCCCGTGACATC 840
Db 260 LysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProValSerIle 279
QY 841 CTGCACATCAAGCAGGCGCCCAAGAGGCCCTTCCGCACTACGTGACCGCTTCTTCAAG 900
Db 280 LeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhePheLys 299
QY 901 ACCCTGGCGCGCGAGAGAGACCAAGAGGTGAAGAACTGGATGACCGACACCTGCTG 960
Db 300 ThrLeuArgAlaGluGlnAlaThrGlnGluValLysAsnTrpMetThrAspThrLeuLeu 319
QY 961 GTGCAGAACGCGCAACCCCGACTGCAGAACCATCTCGCGCGCTCTGCGGCCCGCGCCAGC 1020
Db 320 ValGlnAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeuGlyProGlyAlaThr 339
QY 1021 CTGGAGAGATGATGACCGCTGCAGAGGGCGGTGGCGGCCCGCCAGCCACAGGCCCGCTG 1080
Db 340 LeuGluGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLysAlaArgVal 359
QY 1081 CTGGCGGAGGCGGATGAGCCAGGCCCAACCC--AGCGTGAATGATGCAAGAGCAACTTC 1137
Db 360 LeuAlaGluAlaMetSerGlnThrAsnSerGlyAsnIleMetMetGlnArgSerAsnPhe 379
QY 1138 AAGGCGCGCGCGCGCATGCTCAAGTCTTCAACTGCGGCAAGAGAGGCCACATCGCCCGC 1197
Db 380 LysGlyProArgArgIleValLysCysPheAsnCysGlyLysGluGlyHisIleAlaArg 399
QY 1198 AACTGCGCGCGCGCGCGCAAGAGAGGCTGCTGGAAGTGGGCAAGAGAGGCGCCACAGATG 1257
Db 400 AsnCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGluGlyHisGlnMet 419

QY 1258 AAGGCTGCACCGGCGCCAGGCCAACTTCTGGGCAAGATCTGGCCAGCAGCAAGGGC 1317
Db 420 LysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSerHisLysGly 439
QY 1318 CGCCCGGCAACTCTCTGCAGAGCGCCCGCCAGCCCAAGAGAGAGAGAGAGAGAGAG 1377
Db 440 ArgProGlyAsnThrLeuGlnAsnArgProGluProThrAlaProProAlaGluSerPhe 459
QY 1378 CGCTTCGAGAGAGACCAACCCCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1437
Db 460 ArgPheGluGluThrThrProAlaProLysGlnGluProIleGluArgGluProLeuThr 479
QY 1438 AGCCTGAAGAGCTGTTGCGCAACGACCCCTGAGCCAG 1476
Db 480 SerLeuLysSerLeuPheGlySerAspProLeuSerGln 492

RESULT 9
US-10-332-413-8
; Sequence 8, Application US/10332413
; Publication No. US20040116660A1
; GENERAL INFORMATION:
; APPLICANT: Johnstone, Robert Edward
; APPLICANT: Swanstrom, Ronald Ivar
; APPLICANT: Morris, Lynn
; APPLICANT: Karim, Salim Abdool
; APPLICANT: Williamson, Carolyn
; TITLE OF INVENTION: Process for the Selection of HIV-1 Subtype C Isolates, Selected H
; TITLE OF INVENTION: Subtype Isolates, Their Genes and Modifications and Derivatives T
; FILE REFERENCE: 45669-281993
; CURRENT APPLICATION NUMBER: US/10/332,413
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: PCT/IB01/01208
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,995
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: ZA 2000/3437
; PRIOR FILING DATE: 2000-07-10
; PRIOR APPLICATION NUMBER: ZA 2000/4924
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-332-413-8

Alignment Scores:
Pred. No.: 3,24e-120 Length: 631
Score: 2435.00 Matches: 458
Percent Similarity: 96.96% Conservative: 20
Best Local Similarity: 92.90% Mismatches: 13
Query Match: 87.84% Indels: 2
DB: 16 Gaps: 2

US-09-475-704A-3 (1-1479) x US-10-332-413-8 (1-631)

QY 1 ATGGGCGCGCGCGCCAGCATCTCGCGCGCGCAAGCTGAGCGCTGGAGCGCATCCGC 60
Db 31 MetAlaAlaArgAlaSerIleLeuArgGlyGluLysLeuAspLysTrpGluLysIleArg 50
QY 61 CTGGCGCGCGCGCGCAAGAGTGTCTACATGATGAAGCACCTGTGTGGCGAGCGCGAG 120
Db 51 LeuArgProGlyGlyLysLysHisTyrMetLeuLysHisIleValTrpAlaSerArgGlu 70
QY 121 CTGGAGAAGTTCGCCCTGAACCCCGCGCTGCTGAGAGAGAGAGAGAGAGAGAGAGATC 180
Db 71 LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrSerGluGlyCysLysGlnIle 90
QY 181 ATCCGCGAGCTGCACCCCGCCCTGCAGACCGGCGAGAGAGAGCTGAAGACCTGTTCAAC 240
Db 91 MetLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluGluLeuLysSerLeuTyrAsn 110


```
QY 241 ACCGTGGCCACCTGTACTGCGTGCGAGAGAGATCGCCGACACCAAGGAGGCC 300
Db 111 ThrValAlaThrLeuTyrCysValHisGluIleGluValArgAspThrLysGluAla 130
QY 301 CTGGCAAGATCGAGAGAGAGAGAGAGAGAGAGAGATCGAGAGATCGAGAGGCC 360
Db 131 LeuAspLysIleGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 150
QY 361 GCCGCAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAAACTGACAGGCCAGATG 420
Db 151 AlaAsp---GlyLysValSerGlnAsnTyrProIleValGlnAsnLeuGlnGlyMet 169
QY 421 GTGCACCAAGCCATCAGCCCCCGACCCCTGAAACCGCTGAGGTGATCGAGAGAA 480
Db 170 ValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValIleGluGluLys 189
QY 481 GCCTTCAGCCCCGAGGTGATCCCACTGTTCAACCGCCCTGAGCGAGGCCACCCCCAG 540
Db 190 AlaPheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGln 209
QY 541 GACCTGAACACGATGTTGAACACCGTGGCGGCGCCACAGGCCCGCATGACATGCTGAAG 600
Db 210 AspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLys 229
QY 601 GACACCATCAACGAGAGAGGCCCGCGAGTGGAGACCGCGTGACACCCCGTGACGCGCGCCC 660
Db 230 AspThrIleAsnGluGluAlaAlaGluTrpAspArgLeuHisProValHisAlaGlyPro 249
QY 661 ATGCCCCCGGCGCAGATGCGCGGAGCGCCGCGGCGGACGACATCGCCGCGCACCAAGACC 720
Db 250 IleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerThr 269
QY 721 CTGCAGAGACGATCGCCTGGATGACACGACCAACCCCCCATCCCGCTGGCGGCGCATCTAC 780
Db 270 LeuGlnGluGlnIleAlaTrpMetThrSerAsnProProIleProValGlyAspIleTyr 289
QY 781 AAGCGGTGATCATCTCTGGGCGCTGAACAAGATCGTGGCGATGTACACCCCGTGACATC 840
Db 290 LysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProValSerIle 309
QY 841 CTGCACATCAAGCAGAGGCCCGCAAGAGAGCCCTTCGCGGACTACGTGACCGCTTCTTCAAG 900
Db 310 LeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhePheLys 329
QY 901 ACCCTGGCGCGCGGAGCAGACACCCAGAGAGGTGAAGAACTGGATGACCGACACCTGCTG 960
Db 330 ThrLeuArgAlaGluGlnAlaThrGlnGluValLysAsnTrpMetThrAspThrLeuLeu 349
QY 961 GTGCAGAAAGCCCAACCCCGACTGCAGAACCATCTCTGGCGCTCTCGGCCCCGGCGGACG 1020
Db 350 ValGlnAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeuGlyProGlyAlaThr 369
QY 1021 CTGCAGAGATGATGACCGCCTGCGGAGCGGTGGCGGCGCCACGACCAAGGCCCGGTG 1080
Db 370 LeuGluGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLysAlaArgVal 389
QY 1081 CTGGCGGAGGCGATGAGCCAGGCCACACAC--AGCGTGAATGATGCAGAAAGCAACTTC 1137
Db 390 LeuAlaGluAlaMetSerGlnThrAsnSerGlyAsnIleMetMetGlnArgSerAsnPhe 409
QY 1138 AAGGGCCCCCGGCGCATGTCAGTGCTTCAACTGCGGCAAGAGGGCCACATCGCCCGC 1197
Db 410 LysGlyProArgArgIleValLysCysPheAsnCysGlyLysGluGlyHisIleAlaArg 429
QY 1198 AACTGCGCGCGCCCCCGCAAGAGGGCTGTGAAGTGGCGGCAAGAGGGCCACCATG 1257
Db 430 AsnCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGluGlyHisGlnMet 449
QY 1258 AAGGACTGCACCGAGCGCCAGCCAACTTCTGGGCAAGATCTGGCCGACCCACCAAGGCC 1317
Db 450 LysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSerHisLysGly 469
```

```
QY 1318 CGCCCCGGCAACTTCTTGCAAGACCGCCCCGAGCCACCGCCCCCGCGGAGAGCTTC 1377
Db 470 ArgProGlyAsnPheLeuGlnAsnArgProGluProThrAlaProProAlaGluSerPhe 489
QY 1378 CGCTTCGAGGAGACCAACCCCGCGCCAGAGACGAGAGAGAGAGAGAGAGAGAGAGAG 1437
Db 490 ArgPheGluGluThrThrProAlaProLysGlnGluProIleGluArgGluProLeuThr 509
QY 1438 AGCCTGAAGAGCTGTTCGGCAACGAGACCCCTGAGCCAG 1476
Db 510 SerLeuLysSerLeuPheGlySerAspProLeuSerGln 522

RESULT 10
US-10-339-217-110
; Sequence 110, Application US/10339217
; Publication No. US20030198648A1
; GENERAL INFORMATION:
; APPLICANT: Buechter, Douglas
; APPLICANT: Hou, Xiaohong
; APPLICANT: Marlor, Christopher W.
; APPLICANT: Rice, William G.
; APPLICANT: Yang, Wengang
; TITLE OF INVENTION: Methods for Identifying Compounds Which Inhibit Binding of
; TITLE OF INVENTION: Nucleocapsid 7 Protein to HIV-1 RNA
; FILE REFERENCE: 111021.143(ACH-US1)
; CURRENT APPLICATION NUMBER: US/10/339, 217
; PRIOR FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: US 60/347,369
; PRIOR FILING DATE: 2002-01-11
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: FaSTSeq for Windows Version 4.0
; SEQ ID NO 110
; LENGTH: 492
; TYPE: PR
; ORGANISM: HIV-1
US-10-339-217-110

Alignment Scores:
Pred. No.: 1.37e-119 Length: 492
Score: 2423.00 Matches: 451
Percent Similarity: 96.96% Conservative: 27
Best Local Similarity: 91.48% Mismatches: 13
Query Match: 87.41% Indels: 2
DB: 14 Gaps: 2

US-09-475-704A-3 (1-1479) x US-10-339-217-110 (1-492)
QY 1 ATGGGCGCGCGCGCCAGCATCTCTGCGCGCGGCGGCAAGCTGAGCCCTGGAGCCATCCGC 60
Db 1 MetGlyAlaArgAlaSerIleLeuArgGlyGluLysLeuAspThrTrpGluLysIleArg 20
QY 61 CTGGCGCGCGCGCGCAAGAGTGTCTACATGATGAAGCAACCTGTGTGGGCCACCGCGAG 120
Db 21 LeuArgProGlyGlyLysLysCysTyrMetLeuLysHisIleValTrpAlaSerArgGlu 40
QY 121 CTGAGAAAGTTCGCCCTGAACCCCGCGCTGCTGAGAGACCAAGGAGGGCTGCAAGCAGATC 180
Db 41 LeuGluArgPheSerLeuAsnProGlyLeuLeuGluThrSerGluGlyCysLysGlnIle 60
QY 181 ATCCGCCAGCTGCAACCCCGCGCTGCAAGACCGGCGAGAGAGCTGAAGAGCTGTCAAC 240
Db 61 MetLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluGluLeuLysSerLeuTyrAsn 80
QY 241 ACCGTGGCCACCTGTACTGCGTGACGAGAAATCGAGGTCCGCGACACCAAGAGGCC 300
Db 81 ThrValAlaThrLeuPheCysValHisGluLysIleAlaValaArgAspThrLysGluAla 100
QY 301 CTGCAGAAAGATCGAGAGAGAGAGAGAGAGAGAGAGATCGAGAGATCCAGAGGCCAGGCC 360
Db 101 LeuAspLysIleGluGluGluGluGlnAsnLysSerGlnGlnLysThrGlnGlnAlaLysAla 120
QY 361 GCCGCAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAAGACTGACAGGCCAGATG 420
```

Db 121 Alaasp---GlyThrValSerGlnAsnTyrProIleValGlnAsnLeuGlnGlnMet 139
QY 421 GTGCACCAAGGCCATCAGCGCCCGCACCTGAAAGCGCTGGGTGAAGGTGATCGAGGAGAAG 480
Db 140 ValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValIleGluGluLys 159
QY 481 GCCTTCAGCCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGCGCCACCCCCAG 540
Db 160 AlaPheSerProGluValIleProMetPheThrAlaLeuSerGluGluAlaThrProGln 179
QY 541 GACCTGAACACGATGTTGAACACCGTGGCGGGCCACAGCGCCGCATGACAGATGCTGAAG 600
Db 180 AspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLys 199
QY 601 GACACCATCAACGAGAGGCGCGCGAGGTGGACCGCGTGACACCCCGTGACGCGCGGCC 660
Db 200 AspThrIleAsnGluGlnAlaAlaGluTrpAspArgLeuHisProValHisAlaGlyPro 219
QY 661 ATGCCCCCGCCAGATGCGCGAGCGCGCGCGGAGCGACATCGCGCGCACACCAAGACC 720
Db 220 IleAlaProGluGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerThr 239
QY 721 CTGCAGAGCAGATCGCGCTGGATGACCAACCCCGCATCCCGTGGCGCAGCATCTAC 780
Db 240 LeuGlnGluGlnIleAlaTrpMetThrAsnAsnProProValProValGlyAspIleTyr 259
QY 781 AAGCGGTGATCATCTGGCGCTGAACAGATCGTCCGATGTACAGCCCGTGAGCATC 840
Db 260 LysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProValSerIle 279
QY 841 CTGACATCAAGCAGGCGCCCAAGAGCGCTTCCGCGACTACGTGACCGCTTCTTCAAG 900
Db 280 LeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhePheLys 299
QY 901 ACCCTGCGCGCCGAGCAGACACCAAGAGGTGAAGAAGCTGATGACCGACACCTGCTG 960
Db 300 ThrLeuArgAlaGluGlnAlaThrGlnAspValLysAsnTrpMetThrAspThrLeuLeu 319
QY 961 GTGCAGAACGCCAACCCCGACTGCAGAACCATCTGCGCGCTCTCGCGCCCGCGCGCAGC 1020
Db 320 ValGlnAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeuGlyProGlyAlaThr 339
QY 1021 CTGCAGAGATGATGACCGCGCTGCGCGGGCGGTGGCGGCCCGACCAAGCGCGCTG 1080
Db 340 LeuGluGluMetMetThrAlaCysGlnGlyValGlyGlyProSerHisLysAlaArgVal 359
QY 1081 CTGCGCGAGGCGCATGAGCCAGGCC---AACACACAGCTGATGATGCAAGAACAACTTC 1137
Db 360 LeuAlaGluAlaMetSerGlnThrAsnAsnAlaAsnIleMetMetGlnArgSerAsnPhe 379
QY 1138 AAGGCGCCCCCGCGCATCTGTCAAGTGTCTCAACTGCGGCAAGAGGGCCACATCGCCCGC 1197
Db 380 LysGlyProArgArgIleIleLysCysPheAsnCysGlyLysGluGlyHisLeuAlaArg 399
QY 1198 AACTGCGCGCGCGCGCGCAAGAGGGGTGCTGGAAGTCCGGCAAGAGGGCCACAGATG 1257
Db 400 AsnCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGluGlyHisGlnMet 419
QY 1258 AAGGACTGCACCGAGCGCCAGGCCAACTTCTGCGGCAAGATCTGGCCACGCCAACAGGGC 1317
Db 420 LysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSerHisLysGly 439
QY 1318 CGCCCCCGCAACTTCTGTGCAGAGCGCGCGCGAGCCCGCGCCCGCGCGAGAGGTTTC 1377
Db 440 ArgProGlyAsnPheLeuGlnAsnArgProGluProThrAlaProProAlaGluSerPhe 459
QY 1378 CGCTTCAGAGAGACCAACCCCGCGCGCAAGAGCAGAGCAAGACCGCGAGACCTTGACC 1437
Db 460 ArgPheGluGluThrThrProAlaProLysGlnGluProArgGluArgGluProLeuThr 479
QY 1438 AGCTGAAGAGCGCTGTTGGCAACGACCCCTGAGCCAG 1476
Db 480 SerLeuLysSerLeuPheGlySerAspProLeuSerGln 492

RESULT 11
US-10-339-217-109
; Sequence 109, Application US/10339217
; Publication No. US20030198648A1
; GENERAL INFORMATION:
; APPLICANT: Buechter, Douglas
; APPLICANT: Hou, Xiaohong
; APPLICANT: Marlor, Christopher W.
; APPLICANT: Rice, William G.
; APPLICANT: Yang, Wengang
; TITLE OF INVENTION: Methods for Identifying Compounds Which Inhibit Binding of
; TITLE OF INVENTION: Nucleocapsid 7 Protein to HIV-1 RNA
; FILE REFERENCE: 111021.143(ACH-US1)
; CURRENT APPLICATION NUMBER: US/10/339,217
; PRIOR APPLICATION NUMBER: US 60/347,369
; PRIOR FILING DATE: 2002-01-11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 109
; LENGTH: 492
; TYPE: PRT
; ORGANISM: HIV-1
US-10-339-217-109

Alignment Scores:
Pred. No.: 3,19e-119 Length: 492
Score: 2416.00 Matches: 455
Percent Similarity: 95.74% Conservative: 17
Best Local Similarity: 92.29% Mismatches: 19
Query Match: 87.16% Indels: 2
DB: 14 Gaps: 2

US-09-475-704A-3 (1-1479) x US-10-339-217-109 (1-492)
QY 1 ATGGGCGCGCGCCAGCATCTCTGCGCGCGGCAAGCTGAGCCCTGGAGCGCATCCGC 60
Db 1 MetGlyAlaSerAlaSerIleLeuArgGlyGlyLysLeuAspLysTrpGluLysIleArg 20
QY 61 CTGCGCCCCCGCGCAAGAGTGTACATGATGAAGCAACCTGGTGTGGCCAGCGCGAG 120
Db 21 LeuArgProGlyGlyLysLysLysTyrArgLeuLysHisLeuValTrpAlaSerArgGlu 40
QY 121 CTGGAAGAATTCCGCTGAAACCCCGCGCTGCTGGAAGACCAAGGAGGCTGCAAGCAGATC 180
Db 41 LeuGluArgPheAlaLeuAsnSerGlyLeuLeuGluThrAlaGluGlyCysLysGlnIle 60
QY 181 ATCCGCCAGCTTGACACCCCGCTGCAAGCCGGCAGCGAGAGCTGAAGAGCTTCAAC 240
Db 61 IleLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluGluLeuLysSerLeuTyrAsn 80
QY 241 ACCGTGGCCACCTGTACTGCGGTGCACGAGAAGATCGAGGTCCGCGACCAAGAGGGCC 300
Db 81 ThrValAlaThrLeuTyrCysValHisAlaGlyIleGluValArgAspThrLysGluAla 100
QY 301 CTGGAACAAGATGAGAGAGAGAGCAACAAGTGGCCAGCAGAGATCCAGAGCGCGAGGCC 360
Db 101 LeuAspLysIleGluGluGluGlnAsnLysCysGlnGlnLysThrGlnGlnAlaLysGlu 120
QY 361 GCGGCAAGGGGCAAGGTGAGCGCAACTACCCCATCGTGCAAACTGACAGGCGCAGATG 420
Db 121 Alaasp---GlyLysValSerGlnAsnTyrProIleValGlnAsnLeuGlnGlnMet 139
QY 421 GTGCACCAAGGCCATCAGCCCCCGCACCTGAAAGCGCTGGGTGAAGGTGATCGAGGAGAAG 480
Db 140 ValHisGlnProIleSerProArgThrLeuAsnAlaTrpValLysValIleGluGluLys 159
QY 481 GCCTTCAGCCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGCGCGCACCCCCAG 540
Db 160 AlaPheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGln 179
QY 541 GACCTGAACACGATGTTGAACACCGTGGCGGCCACAGGCGCCCATGACAGTGTGAAG 600

```
|||||
Db 180 AspleuAsnThrMetLeuAsnThrValGlyGlnIleAlaIleMetGlnMetLeuLys 199
QY 601 GACACCATCAACGAGAGGCGCGCGATGGAGCCGCGGTGCACCCCGGTGCAGCGCGCGCC 660
Db 200 AspThrIleAsnGlnIleValIleAlaGlnIleProAspArgLeuHisProValHisIleAlaGlyPro 219
QY 661 ATCGCCCCCGGCGAGATGCGCGAGCGCGCGCGGAGGACATCGCGCGGACACCGAGCACC 720
Db 220 ValAlaProGlyGlnMetArgGlnProArgGlySerAspIleAlaGlyThrThrSerThr 239
QY 721 CTGCAGGAGCAGATCGCTGGATGACGACCAACCCCGCATCCCGGTGGCGGACATCTAC 780
Db 240 LeuGlnGlnGlnIleThrThrMetThrSerAsnProProValProValGlyAspIleTyr 259
QY 781 AAGCGGTGATCATCTGGGCTGAAACAAGATCGTGGGATGTACAGCCCGGTGAGCATC 840
Db 260 LysArgTyrIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProValSerIle 279
QY 841 CTGCATCAAGCAGGCGCGCAAGGAGCGCTTCGCGGACTACGTTGACCGCTTCTTCAAG 900
Db 280 LeuAspIleLysGlnGlyProLysGlnProPheArgAspTyrValAspArgPhePheLys 299
QY 901 ACCCTGGCGCGCGAGCAGACCCGAGAGGTGAAGAAGTGTGATGACGACCCGCTGCTG 960
Db 300 ValLeuArgAlaGlnIleAlaThrGlnAspValLysAsnTyrMetThrAspThrLeuLeu 319
QY 961 GTGCAGAACGCCAACCGGACTGCAAGACCATCTCGCGGCTCTCGCGCGCGCGCGAGC 1020
Db 320 ValGlnAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeuGlyProAlaIleSer 339
QY 1021 CTGCAGAGATGATGACCGCTGCGGCGGCGGTGCGCGCGCGCGCGAGCGCGGTG 1080
Db 340 LeuGlnGlnMetMetThrAlaCysGlnGlyValGlyGlyProSerHisLysAlaArgVal 359
QY 1081 CTGGCGGAGCGGATGAGCGGCGGCAAC--ACCAGCGTGTGATGAGAGAAGCAACTTC 1137
Db 360 LeuAlaGlnAlaMetSerGlnAlaAsnThrThrAsnIleMetMetGlnLysSerAsnPhe 379
QY 1138 AAGGCGCGCGCGCGCATCGTCAAGTGTTCACCTGCGGCAAGAGGCGCACATCGCGCGC 1197
Db 380 LysGlyProArgTyrGlnValLysCysPheAsnCysGlyLysGlnGlnHisIleAlaLys 399
QY 1198 AACTGCGCGCGCGCGCGCAAGAGGCGTGTGAACTGCGGCAAGAGGCGCGGCGAGATG 1257
Db 400 AsnCysArgAlaProArgLysLysGlyCysTyrLysCysGlyLysGlnGlnHisIleGlnMet 419
QY 1258 AAGCACTGCACCGAGCGCGGCGCAACTTCCTGGGCAAGATCTGGCGCGGCGGCGGCGC 1317
Db 420 LysAspCysThrGlnArgGlnAlaAsnPheLeuGlyLysIleTyrProSerHisLysGly 439
QY 1318 CGCGCGCGCAACTTCCTGCAGAGCGCGCGCGGCGGCGGCGCGCGCGCGGCGGAGCTTC 1377
Db 440 ArgProGlyAsnPheLeuGlnSerArgProGlnProThrAlaProProAlaGlnSerPhe 459
QY 1378 CGCTTCGAGGAGACCAACCGCGCGGCGCAAGAAGCAGAGCAAGCAAGCAAGCAAGCAAGC 1437
Db 460 ArgPheGlnGlnThrThrProAlaProLysGlnGlnProLysAspArgGlnProLeuThr 479
QY 1438 AGCCTGAAGAGCCTGTTCCGCAACGACCGCGCTGAGCGAG 1476
Db 480 SerLeuLysSerLeuPheGlySerAspProLeuSerGln 492
```

RESULT 12
US-10-339-217-140

; Sequence 140, Application US/10339217

; Publication No. US20030198648A1

; GENERAL INFORMATION:

; APPLICANT: Buechter, Douglas

; APPLICANT: Hou, Xiaohong

; APPLICANT: Marlor, Christopher W.

; APPLICANT: Rice, William G.

; APPLICANT: Yang, Wengang

```
; TITLE OF INVENTION: Methods for Identifying Compounds Which Inhibit Binding of
; TITLE OF INVENTION: Nucleocapsid 7 Protein to HIV-1 RNA
; FILE REFERENCE: 111021.143(ACH-US1)
; CURRENT APPLICATION NUMBER: US/10/339,217
; CURRENT FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: US 60/347,369
; PRIOR FILING DATE: 2002-01-11
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 140
; LENGTH: 491
; TYPE: PRT
; ORGANISM: HIV-1
US-10-339-217-140
```

Alignment Scores:

Pred. No.:	3,39e-119	Length:	491
Score:	2415.50	Matches:	450
Percent Similarity:	96.75%	Conservative:	26
Best Local Similarity:	91.46%	Mismatches:	15
Query Match:	87.14%	Indels:	1
DB:	14	Gaps:	1

US-09-475-704A-3 (1-1479) x US-10-339-217-140 (1-491)

```
QY 1 ATGGGCGCGCGCGCGCGCATCTGCGCGGCGGCAAGTGAAGCCTGGAGCGCATCCGC 60
Db 1 MetGlyAlaArgAlaSerIleLeuArgGlyGlyLysLeuAspLysTyrGlyLysIleArg 20
QY 61 CTGGCGCGCGCGCGCAAGAAGTGTACATGATGAAGCACCTGTGTGGCGCGCGCGAG 120
Db 21 LeuArgProGlyGlyLysLysHisTyrMetLeuLysHisIleValTyrAlaSerArgGlu 40
QY 121 CTGCAGAAATTGCGCTGAACCGCGCGCTGTGAAGACGAGCGAGCGAGCGGTGAGATC 180
Db 41 LeuGlnArgPheAlaValAsnProGlyLeuLeuGlnThrAlaGlnGlyCysLysGlnIle 60
QY 181 ATCGCGCAGCTGCACCCCGCGCTGCAGACCGCGGCGAGCGAGGAGTGAAGACTGTTCAAC 240
Db 61 IleLysGlnLeuGlnProAlaLeuGlnThrGlyThrGlnGlnLeuArgSerLeuTyrAsn 80
QY 241 ACCGTGGCGCACCTGTACTGCGTGCACGAGAATCGAGGTCCGCGACCAAGAGGCC 300
Db 81 ThrValAlaThrLeuTyrCysValHisIleAlaGlnIleGlnValArgAspThrLysGlnAla 100
QY 301 CTGCAGCAAGATCGAGGAGGAGGAGCAACAAGTGCAGGACGAGCAAGATCCAGCGGAGGCC 360
Db 101 LeuAspArgIleGlnGlnGlnGlnAsnLysSerGlnGlnLysThrGlnGlnAlaAsnGlu 120
QY 361 GCGGACAAAGGCGAAGGTGAGCGGCAACTACCCCATCGTGCAAGAAGCTGCAGGCGAGATG 420
Db 121 AlaAsp--GlyLysValSerGlnAsnTyrProIleValGlnAsnLeuGlnGlnMet 139
QY 421 GTGCACGAGCGCATCAGCGCGCGGCGACCGCTGAAGCGCTGGTGAAGGTGATCGAGAGAAG 480
Db 140 ValHisGlnAlaIleSerProArgThrLeuAsnAlaTyrValLysValIleGlnGlnLys 159
QY 481 GCCTTCAGCGCGCGAGGTGATCCCATGTTCAACCGCGCTGAGCGAGGCGCGGCGGCGGAG 540
Db 160 AlaPheSerProGlnValIleProMetPheThrAlaLeuSerGlnGlyAlaThrProGln 179
QY 541 GACTGAACAACGATGTTGAACACCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAG 600
Db 180 AspleuAsnThrMetLeuAsnThrValGlyGlnHisGlnAlaIleMetGlnMetLeuLys 199
QY 601 GACACCATCAACGAGAGGCGCGCGGAGTGGACCGCGGTGCACCCCGGTGCAGCGCGCGCC 660
Db 200 AspThrIleAsnGlnGlnAlaIleAlaGlnIleTyrAspArgLeuHisProValHisIleAlaGlyPro 219
QY 661 ATCGCCCCCGGCGAGATGCGCGGAGCGCGCGGCGGCGGAGCATCGCGCGGCGGAGCACC 720
Db 220 IleAlaProGlyGlnMetArgGlnProArgGlySerAspIleAlaGlyThrThrSerSer 239
```

```

QY 721 CTGCAGAGAGCAGATCGCCTGGATGACCAGCAACCCCCCATCCCGCTGGGCGCATCTAC 780
    |||||
Db 240 LeuGInGluGlnIleAlaTyrMetThrGlyAsnProProValGlyAspIleTyr 259
QY 781 AAGCGGTGATCATCTGGGCGCTGAAACAAGATCGTGGGATGTACAGCCCCGTGAGCATC 840
    |||||
Db 260 LysArgTyrIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProValSerIle 279
QY 841 CTGCACATCAAGCAGGGGCCCCCAAGAGCCCCCTTCGGCGACTACGTGACCGGCTTCTCAAG 900
    |||||
Db 280 LeuAspIleLeuGlnGlyProLysGluProPheArgAspTyrValAspArgPhePheLys 299
QY 901 ACCCTGCGCGCGAGCAGACAGCACCCAGAGGTGAAGACTGATGACCGACACCTGCTG 960
    |||||
Db 300 ThrLeuArgAlaGluGlnAlaThrGlnAspValLysAsnTyrMetThrAspThrLeuLeu 319
QY 961 GTGCAGAACCGCAACCCCGACTGCAAGACCATCTGCGGCTCTGGCCCCGGCGCCAGC 1020
    |||||
Db 320 ValGlnAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeuGlyProGlyAlaSer 339
QY 1021 CTGGAGAGATGATGACCGCCTGCGCAGGGCGGTGGGCGGCCAGCCACAAGCGCGGTG 1080
    |||||
Db 340 LeuGluGluMetMetThrAlaCysGlnGlyValGlyGlyProSerHisLysAlaArgVal 359
QY 1081 CTGGCCGAGGCGCATGAGCCAGGCCAACACACCGGTGATGATGACGAAGAAGCAATTCAAG 1140
    |||||
Db 360 LeuAlaGluAlaMetSerGlnThrAsnSerThrIleLeuMetGlnArgSerAsnPheLys 379
QY 1141 GGCCCCCGCGCATCGTCAAGTGCTTCAACTGCGGCAAGAGAGGACCATCGCCGCAAC 1200
    |||||
Db 380 GlyProLysArgIleValLysCysPheAsnCysGlyLysGluGlnHisIleAlaLysAsn 399
QY 1201 TGCCGCGCCCCCGCAAGAAGGGCTGCTGAAGTGCGGCAAGAGGGCCACCATGAAG 1260
    |||||
Db 400 CysArgAlaProArgLysGlyCysTyrLysGlyLysGluGlnHisGlnMetLys 419
QY 1261 GACTGCACCGAGCGCCAGGCCCACTTCTGGGCAAGATCTGGCCAGCCACAAGGGCGCG 1320
    |||||
Db 420 AspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTyrProSerHisLysGlyArg 439
QY 1321 CCCGCAACTTCTGTCAGAGCGCGCCGAGGCCACCGCCCCCGCGAGAGCTTCCGC 1380
    |||||
Db 440 ProGlyAsnPheLeuGlnSerArgProGluProThrAlaProProAlaGluSerPheArg 459
QY 1381 TTCGAGAGAGACCAACCCCGCGCAAGAAGAGAGAGCAAGAGACCGAGACCTGACACGC 1440
    |||||
Db 460 PheGluGluThrThrProAlaLeuGlnGlnGlyProLysAspArgGluProLeuThrSer 479
QY 1441 CTGAAGAGCCCTGTCGCGCAACGACCCCTGAGCCAG 1476
    |||||
Db 480 LeuArgSerLeuPheGlySerAspProLeuSerGln 491

RESULT 13
US-10-339-217-104
: Sequence 104, Application US/10339217
: Publication No. US20030198648A1
: GENERAL INFORMATION:
: APPLICANT: Buechter, Douglas
: APPLICANT: Hou, Xiaohong
: APPLICANT: Marlor, Christopher W.
: APPLICANT: Rice, William G.
: APPLICANT: Yang, Wengang
: TITLE OF INVENTION: Methods for identifying Compounds which Inhibit Binding of
: FILE REFERENCE: 111021.143(ACH-US1)
: CURRENT APPLICATION NUMBER: US/10/339,217
: PRIOR FILING DATE: 2003-01-09
: PRIOR APPLICATION NUMBER: US 60/347,369
: NUMBER OF SEQ ID NOS: 144
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 104
: LENGTH: 491

```

:	TYPE:	PRT
:	ORGANISM:	HIV-1
US-10-339-217-104		
Alignment Scores:		
Pred. No.:	8.95e-119	Length: 491
Score:	2407.50	Matches: 449
Percent Similarity:	96.34%	Conservative: 25
Best Local Similarity:	91.26%	Mismatches: 17
Query Match:	86.85%	Indels: 1
DB:	14	Gaps: 1
US-09-475-704A-3 (1-1479) x US-10-339-217-104 (1-491)		
QY	1 ATGGCGCGCCGCGGCATCTTCGCGCGCGCAAGTGAACGCCCTGGAGCGCATCCG	60
Db	1 MetcIyAlaArgAlaSerlLeuValrgGlulysleuAspLysTrpGlulysileArg	20
QY	61 CTGCGCCCCGCGCGCAAGAAGTGCTACATGATGAAGCCTGTGTGGCCAGCCGCGAG	120
Db	21 LeuArgrProgllyArglysHieTyMetleuysHieValrPalaserArglu	40
QY	121 CTGGAAGAAGTTGCGCCTGAACCCCGCCTGTGAGACACGAGGCGTCAAGCAGATC	180
Db	41 LeugluArgPheAlaleuAsnProglyleuleugluThrserInglyCylsglnlie	60
QY	181 ATCCGCGAGCTGCACCCCGCCTGCAGACCGGACGAGGAGAGCTGTAAGCCTGTTCAAC	240
Db	61 IleysglnleuHisproAlaleuLysThrglyThrglugluLeuArgserleutyrsn	80
QY	241 ACCGTGGCCACCCTGTACTGCGTGACAGAGAAGTAGAGTCCGCGCACCAAGAGGCC	300
Db	81 ThrValAlatThrleutyrcysvalHisglnasnllegluValArgAspThrLysgluaa	100
QY	301 CTGGAACAAGATCGAGAGGAGAGCAGAACAAGTGCCACGAGAAGATCCAGAGGCCGAGGCC	360
Db	101 LeuAspLysileglugluglinslnlysserglnglnlysthrnglnalalyala	120
QY	361 GCCGCAAGGCGCAAGGTGAGCCAGAACTACCCCATTCTGCAGAACCTGCAGGGCCAGATG	420
Db	121 Alaaspglugly--ValserglnasnTyrrProileValglnasnleunglnglnmet	139
QY	421 GTGCACCGAGCCATCAGCCCCCGCAACCTGAACGCTGGTGAAAGGTGATCGAGAGAG	480
Db	140 ValHisglnAlaIleserProArgrThrleuAsnAlatrPallysValilegluglulys	159
QY	481 GCCTTCAGCCCCGAGGTGATCCCAATGTTCAACGCGCTGACGAGGCGCCACCCCCAG	540
Db	160 AlapheserProgluValillePrometPhethrAlaleuSerlglyAlatrhrProgh	179
QY	541 GACCTGAACACGATGTTGAACACCGGTGGGCGGCACACGAGCCGCGCATGAGTGTGAAG	600
Db	180 AspLeuAsnThrMetleuAsnThrValglglyHisglnAlaAlameGlnMetleuys	199
QY	601 GACACCATCAAAGAGAGGCGCGCGGTGGAGACCGCGTGCACCCCGTGCACGCGCGCCC	660
Db	200 AsprthrIleasnlglnAlaAlagluTrpasArgleuHisproValHisAlaglyPro	219
QY	661 ATGCCCCCGCGCAGATGCGGAGCGCGCGCGGACGAGCATCGCCGCGCACCAAGCAGACC	720
Db	220 AlaAlaProglylnMetArggluProArgrglyserAspIleAlaglyThrThrserThr	239
QY	721 CTGCAAGAGCAGATCGCCTGATGACCAACCCCCCATCCCGGTGGGCGACATCTAC	780
Db	240 LeuglngluglnlieAlatrmetThrglyAsnProProvalProvalglyAspiletyr	259
QY	781 AAGCGGTGATCATCTCTGGGCTGTGAACAAGATCGTGGGATGTACAGCCCGTGAGCATC	840
Db	260 LysArgrTrpilelleuglyleuAsnlysillevAlargMettyrserProvalserilie	279
QY	841 CTGCAATCAAGCAGGCGCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTTCAAG	900
Db	280 LeuAspilelysglnglyProlyeglnupropheargsAsptyrValabArgrphephelys	299

```

QY      901 ACCCTGCGCGCCGAGACAGCACCCAGAGGTGAAGA CTGATGACCGACACCCCTGCTG 960
      |||||||
Db      300 ValLeuArGAlaGlInAlaThrGlnAsnProValLysAsnTrpMetThrAspThrLeuLeu 319
QY      961 GTGCAGAACGCCAACCCCGCACTGCAAGACCATCTCGCGCTCTCGGCCCGCGCGCAGC 1020
      ::::|
Db      320 ILegInAsnAlaAsnProAspCysLysThrILeuleuValaLeuGlyProAlaAlaSer 339
QY      1021 CTGAGAGAGATGATGACCGCTGCGCAGGGCGTGGCGGCCCCAGCCACAAGGCCCGCGTG 1080
      |||||||
Db      340 LeuGlnGlnMetMetThrAlaCysGlnGlnValGlyGlyProGlyHisLysAlaArgVal 359
QY      1081 CTGCGCCGAGCGGATGAGCCAGGCCAACACACGCGTGATGATGACAGAAAGCAACTTCAAG 1140
      |||||||
Db      360 LeuAlaGlnAlaMetSerGlnAlaAsnSerAsnILEuMetMetGlnArgSerAsnRheLys 379
QY      1141 GGGCCCGCGCGCATCTCAAGTGTCTTCACTGCGGCAAGAGGCGCCACATCGCCCGCAAC 1200
      ||| ::::|
Db      380 GlySerLysArgILEValLysCysRheAsnLysGlyLysGlnGlnHisILEAlaArgAsn 399
QY      1201 TGCCCGCGCCCCCGCAAGAGGCTGTGAGTGTGAGTGTGCGCAAGAGGCGCCACAGATGAAG 1260
      |||||||
Db      400 CysArgAlaProArgLysLysGlyCysTrpLysCysGlyGlnGlnLysILELysMetLys 419
QY      1261 GACTGCACCGAGCGCCAGGCCCAACTTCTTGCGCAAGATCTGCGCCAGCCACAAGGCGCGC 1320
      |||||||
Db      420 AspCysThrGlnArgGlnAlaAsnRheLeuGlyLysILETrpProSerHisLysGlyArg 439
QY      1321 CCGCGCACTTCTCTGCAAGCGCGCGCCAGGCCACCGCCCCCGCGAGAGCTTCCGC 1380
      |||||||
Db      440 ProGlyAsnRheLeuGlnSerArgProGlnProThrAlaProProAlaGlnUserRheArg 459
QY      1381 TTGAGAGAGACCAACCCCGCGCCAGAAAGCAGAGAGAGAGAGACCGCGCAGACCTGACGAC 1440
      |||||||
Db      460 PheGlnGlnThrThrProAlaProLysGlnGlnLysSerLysAspArgGlnProLeuILESer 479
QY      1441 CTGAAGAGCCTGTTGCGCAACGACCCCTGAGCCAG 1476
      |||||||
Db      480 LeuLysSerLeuPheGlySerAspProSerSerGln 491

```

```

RESULT 14
US-10-339-217-143
; Sequence 143, Application US/10339217
; Publication No. US20030198648A1
; GENERAL INFORMATION:
; APPLICANT: Buechter, Douglas
; APPLICANT: Hou, Xiaohong
; APPLICANT: Marlor, Christopher W.
; APPLICANT: Rice, William G.
; APPLICANT: Yang, Wengang
; TITLE OF INVENTION: Methods for Identifying Compounds Which Inhibit Binding of
; FILE REFERENCE: Nucleocapsid 7 Protein to HIV-1 RNA
; CURRENT APPLICATION NUMBER: US/10/339,217
; PRIOR APPLICATION NUMBER: 2003-01-09
; PRIOR FILING DATE: 2002-01-11
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 143
; LENGTH: 491
; TYPE: PRT
; ORGANISM: HIV-1
US-10-339-217-143

```

```

Alignment Scores:
Pred. No.:      3.39e-118      Length:      491
Score:          2396.50       Matches:     449
Percent Similarity: 95.53%    Conservative: 21
Best Local Similarity: 91.26%  Mismatches:    21
Query Match:     86.45%      Indels:       1
DB:              14          Gaps:         1

```

```

US-09-475-704A-3 (1-1479) x US-10-339-217-143 (1-491)
QY      1 ATGGGCGCGCGCCAGCATCTGCGCGCGCGCAAGCTGACGCTGGAGCGCATCCGC 60
      |||||||
Db      1 MetGlyAlaArgAlaSerILEuArgGlyGlyLysLeuAspLysTrpGlnLysILEArg 20
QY      61 CTGCGCGCGCGCGCAAGAGTGTACATGATGAAGCACCTGTGTGGCGCGCAGC 120
      |||||||
Db      21 LeuArgProGlyLysLysHisTrpMetILEYHisLysValTrpAlaSerArgGln 40
QY      121 CTGAGAGAGTTGCGCTGAACCCCGCGCTGTGAGAGACGAGCGGCTGCAAGCATC 180
      |||||||
Db      41 LeuGlnArgRheAlaLeuAsnProGlyLeuLeuGlnThrSerGlnGlyCysLysGlnILE 60
QY      181 ATCCGCGAGCTGACCCCGCGCTGCAGACCGCGCAGAGAGCTGAAGAGCTGTCAAC 240
      ||| ::::|
Db      61 ILELysGlnLeuGlnProAlaLeuGlnThrGlyThrGlnGlnLeuArgSerLeuHisAsn 80
QY      241 ACCGTGCGCACCTGTACTGTGTCGTCAGAGAGATCGAGTGTGCGGACACCAAGAGGCC 300
      |||||||
Db      81 ThrValAlaThrLeuTyrcysValHisValaGlyILELysILEArgAspThrLysGlnAla 100
QY      301 CTGCACAAGATGAGAGAGAGACAGAACAGTGTCCAGAGAGATCCAGAGCGCGAGGCC 360
      |||||||
Db      101 LeuAspLysILEglnGlnGlnGlnLysSerGlnGlnLysThrGlnGlnAlaLysGln 120
QY      361 GCCGACAAGGCGAAGTGAGCCAGAACTACCCCATCGTGCAAGACCTGACAGGCGCAGATG 420
      |||||||
Db      121 AlaAsp--GlyLysValSerGlnAsnTrpProILEValGlnAsnLeuGlnGlnMet 139
QY      421 GTGCACAGGCGCATCAGCCCGCGCACCTGTGAACGCTGGGTGAAGTGATCGAGAGAG 480
      |||||||
Db      140 ValHisGlnAlaLeuSerProArgThrLeuAsnAlaTrpValILEglnGlnLys 159
QY      481 GCCTTCAGCCCGAGGTGATCCCATGTTTACCGCGCTGACGCGAGCGGCCACCCCGCAG 540
      |||||||
Db      160 AlaPheSerProGlnILELeuPheMetPheThrAlaLeuSerGlnGlyAlaThrProGln 179
QY      541 GACCTGAACAGATGTTGAACACCGTGCGCGCCACGAGCGCCCATGACAGATGCTGAAG 600
      |||||||
Db      180 AspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLys 199
QY      601 GACACCATCAAGAGAGAGCGCGCGAGTGGACCGCGTGCAACCCCGTGACGCGCGCCCC 660
      |||||||
Db      200 AspThrILEaGlnGlnAlaGlnAlaGlnTrpAspArgLeuHisProAlaGlnAlaGlyPro 219
QY      661 ATCGCCCCCGCGCAGATGCGCGAGCCCCCGCGCAGCGACATGCGCGCGCACACGACACC 720
      |||||||
Db      220 ILEAlaProGlyGlnMetArgGlnProArgGlySerAspILEAlaGlyThrThrSerThr 239
QY      721 CTGCAGAGAGATGCGCTGTGATGACCAACCCCCCATCCCGCTGGGCGACATCTAC 780
      |||||||
Db      240 LeuGlnGlnGlnILEAlaTrpMetThrGlyAsnProProValProValGlyLysILETyx 259
QY      781 AAGCGGTGATCATCTGTGGCGCTGAACAAGATCGTGGGATGTACAGCCCGTGAGCATC 840
      |||||||
Db      260 LysArgTrpILELeuLeuGlyLeuAsnLysILEValArgMetTySerProValSerILE 279
QY      841 CTGCACATCAAGCAGAGGCGCCCAAGAGCCCTTCCGCGACTACGTGAGACCGCTTCTTAAG 900
      |||||||
Db      280 LeuAspILELysGlnGlyProLysGlnProPheArgAspTyValaAspArgPhePheLys 299
QY      901 ACCCTGCGCGCGCAGACAGACACCCAGAGGTGAAGAACTGTGATGACCGCACACCTGCTG 960
      |||||||
Db      300 ThrLeuArgAlaGlnGlnAlaThrGlnAspValLysAsnTrpMetThrAspThrLeuLeu 319
QY      961 GTGCAGAAAGCCCAACCCCGACTGCAAGACCATCTTGGCGCTCTGGGCGCGCGCAGC 1020
      |||||||
Db      320 ValGlnAsnAlaAsnProAspCysLysThrILELeuArgAlaLeuGlyProGlyAlaSer 339
QY      1021 CTGAGAGAGATGATGACCGCTGCGCAGGCGCGTGGCGGCCCGCCAGCCACAAGCGCGCGTG 1080
      |||||||

```


Db 340 LeuGIuGIuMeMetThrAlaCysGInGIyValGIyGIyProSerHisLysAlaArgVal 359
QY 1081 CTGGCCGAGGGCGATGATGACCGCAACACCGGTGATGATGACAGAACTTCAAG 1140
Db 360 LeuAlaGIuAlaMetSerGIuThrAsnAsnSerIleuMetGIuArgSerAsnPheLys 379
QY 1141 GGGCCCCGGCGATCGTCAAGTGTCTCACTGCGGCAAGAGGGCCACATGCCCCGAAC 1200
Db 380 GIyPheLysArgThrValLysCysPheAsnCysGILySGLySGLyHisIleAlaArgAsn 399
QY 1201 TGCCCGCCCCCGCAAGAGGGCTGTGAGTGTGAGTGTGAGGAGGGCCACAGATGAG 1260
Db 400 CysArgAlaProArgLysLysGIyCysGTTrLysCysGILyLysGILyHisGILyMetLys 419
QY 1261 GACTGACCGGAGCGCCGCAACTTCTGGGCAAGATCTGGCCSAGCCACAGGGCGCC 1320
Db 420 AspCysThrGIuArgGIuAlaAsnPheLeuGIyLysIleTrpProSerHisLysGIyArg 439
QY 1321 CCGGCACTTCTGTGAGAGCGCGCCGAGCCGACCGCCCCCGCGAGAGCTTCCGC 1380
Db 440 ProGIyAsnPheLeuGIuAsnArgProGIuProThrAlaProProAlaGIuSerPheArg 459
QY 1381 TTGAGAGAGACCAACCGCGCCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 460 PheGIuGIuThrThrThrProAlaLeuLysGInGIuGIuLysAspArgGIuProLeuThrSer 479
QY 1441 CTGAAGAGCTGTTCGCAACGACCCCTGAGCCAG 1476
Db 480 LeuLysSerLeuPheGIySerAspProLeuSerGIu 491

RESULT 15
US-10-339-217-108
; Sequence 108, Application US/10339217
; Publication No. US20030198648A1
; GENERAL INFORMATION:
; APPLICANT: Buechter, Douglas
; APPLICANT: Hou, Xiaohong
; APPLICANT: Marlor, Christopher W.
; APPLICANT: Rice, William G.
; APPLICANT: Yang, Wengang
; TITLE OF INVENTION: Methods for Identifying Compounds which Inhibit Binding of
; FILE REFERENCE: 111021.143(ACH-US1)
; CURRENT APPLICATION NUMBER: US/10/339,217
; PRIOR FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: US 60/347,369
; PRIOR FILING DATE: 2002-01-11
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 108
; LENGTH: 496
; TYPE: PRT
; ORGANISM: HIV-1
US-10-339-217-108

Alignment Scores:
Pred. No.: 1.97e-117 Length: 496
Score: 2382.00 Matches: 446
Percent Similarity: 94.77% Conservative: 25
Best Local Similarity: 89.74% Mismatches: 20
Query Match: 85.93% Indels: 6
DB: 14 Gaps: 2

US-09-475-704A-3 (1-1479) x US-10-339-217-108 (1-496)

QY 1 ATGGGCGCCGCGCGCAGCATCTCTGCGCGCGGCAAGCTGAGCGCTGGAGCGCATCCGC 60
Db 1 MetGIyAlaArgAlaSerIleLeuArgGIyGIyLysLeuAspLysTrpGIuLysIleArg 20
QY 61 CTGCGCCCGCGCGGCAAGAGTGTCTCATGATGAAGCACTGTGTGGCCAGCGCGAG 120
Db 21 LeuArgProGIyGIyLysLysLysTyMetLeuLysHisIleuValTrpAlaSerArgGIu 40

QY 121 CTGAGAAAGTTCCGCTGAACCCCGGCTGTGAGAGACCGAGCGGTGACAGCATC 180
Db 41 LeuGIuArgPheAlaLeuAsnSerGIyLeuLeuGIuThrSerAspGIyCysLysGInIle 60
QY 181 ATCCGCGAGTGTGACCCCGGCTGTGACAGCCGCGAGAGAGAGCTGAAGAGCTGTCAAC 240
Db 61 IleGIuGIuLeuGIuProAlaLeuLysThrGIyThrGIuGIuLeuArgSerIleuTyArgn 80
QY 241 ACCGTGGCCACCTGTACTGCGTGCACGAGAAATGAGATGAGTCCCGGACACCAAGAGGCC 300
Db 81 ThrValAlaThrLeuTyTrCysValHisAsnAsnIleGIuIleArgAspThrLysGIuAla 100
QY 301 CTGGACAAGATCGAGAGAGAGACAAGTGC-----CAGCAAGAATC 345
Db 101 LeuAspArgIleGIuGIuGIuGIuLysLysCysGInGIuLysThrGInGIuLysThr 120
QY 346 CAGCAGCCGAGGCGCGCGCAAGGCGCAAGGTAGCCAGAACTTACCCCATCGTGCAAGAC 405
Db 121 GIuGIuValGIuAlaAlaAsp---GIyLysValSerGIuAsnTyrrProIleValGIuAsn 139
QY 406 CTGAGAGGCGCAGATGTGTGACCAAGCCCATCAGCCCCCGCACCTGTGAACGCTGTGAG 465
Db 140 LeuGIuGIuGIuMetValHisGInSerLeuSerProArgThrLeuAsnAlaTrpValLys 159
QY 466 GTGATCGAGAGAGAGGCTTCAAGCCCGGAGGTATCCCATGTTCACCGCCTGAGCGAG 525
Db 160 ValIleGIuGIuLysAlaPheSerProGIuIleIleProMetPheThrAlaLeuSerGIu 179
QY 526 GGGCCCAACCCCGAGGACCTGTGAACACGATGTGAACACCGTGGCGGCCACAGCGCCGC 585
Db 180 GIyAlaThrProGIuAspLeuAsnThrMetLeuAsnThrValGIyHisGInAlaAla 199
QY 586 ATGCAGATGTGAAGGACACCATCAACGAGAGAGCGCGGAGTGGAGCCGCGTGCACCCC 645
Db 200 MetGIuMetLeuLysAspThrIleAsnGIuGIuAlaAlaGIuTrpAspArgLeuHisPro 219
QY 646 GTGCACGCGCGGCGCATCGCCCCCGGCAAGATGCGGAGCGCGCGGAGCGACATCGCC 705
Db 220 ValHisAlaGIyProValAlaProGIyGInMetArgGIuProArgGIySerAspIleAla 239
QY 706 GGCACACACGACACCTGTGAGAGACAGATCGCTGTGATGACCAACACCCCATCCCC 765
Db 240 GIyThrThrSerAsnLeuGIuGIuGIuIleAsnTrpMetThrAlaAsnProIlePro 259
QY 766 GTGGGCGACATCTAACAGCGGTGTGATCATCTGTGGGCTGTGAACAAGATCGTGGAGTGTAC 825
Db 260 ValGIyAspIleTyTrLysArgTrpIleIleLeuGIyLeuAsnLysIleValaArgMetTy 279
QY 826 AGCCCCGTGAGCATCTGTGACATCAAGCAGAGGCCCCCAAGAGACCTTCGCGACTACGTG 885
Db 280 SerProValSerIleLeuAspIleLysGInGIyProLysGIuProPheArgAspTyrrVal 299
QY 886 GACCGCTTCTTCAAGACCTGTGGCGCGGAGAGAGACCAAGAGGTGAAGACTGGATG 945
Db 300 AspArgPhePheLysThrLeuArgAlaGIuGIuAlaThrGIuAspValLysAsnTrpMet 319
QY 946 ACCGACACCTGTGTGTGAGAAAGCCCAACCCGCACTGTCAAGACCATCTGTGGCTCTC 1005
Db 320 ThrAspThrLeuLeuValGIuAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeu 339
QY 1006 GGGCCCCGGCGGAGCTGTGAGAGATGATGACCGCTGTCCAGAGGCGTGGCGGCCCCAGC 1065
Db 340 GIyProGIyAlaThrLeuGIuGIuMetMetThrAlaCysGInGIyValGIyGIyProSer 359
QY 1066 CACAAGGCGCGGTGTGCGGAGGCGATGAGCCAGACCAACAGCGTGATGATGACAG 1125
Db 360 HisLysAlaArgValLeuAlaGIuAlaMetSerGIuThrAsnSerAsnIleMetMetGIu 379
QY 1126 AAGAGCAACTTCAAGGGCGCGCGCGCATGTCAAGTGTCTTCAACTGCGGCAAGAGGCGC 1185
Db 380 AsnSerAsnPheLysGILySerArgArgIleValLysCysPheAsnCysGIyLysValGIy 399
QY 1186 CACATGCGCCGCAACTGCGCGCGCCCCCGCAAGAAAGGCGTGTGGAAGTGGCGCAAGAG 1245

Db	400	HisIleAlaArgAsnCyseArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGlu	419
QY	1246	GGCCACGAGATGAAGGACTGCACCGGAGCGCCAACTTCCTGGGCAAGATCTGGCCC	1305
Db	420	GlyHisGlnMetLysAspCysThrGluArgGlnAlaAsnPhelLeuGlyArgIleTrpPro	439
QY	1306	AGCCACAAGGGCGCGCCGCAACTTCCTGCAGAGCGCCCGAGCCCAACGCCCCCCC	1365
Db	440	SerHisLysGlyArgProGlyAsnPhelLeuGlnSerArgProGluProThrAlaProPro	459
QY	1366	GCCGAGAGCTTCCGCTTCGAGAGACCAACCCCGGCCAGAAGCAGAGACCAAGACCGC	1425
Db	460	AlaGlnSerPheArgPheGluGlnIleThrProValProLysGlnGluProLysAspArg	479
QY	1426	GAGACCTTGACCAAGCTGTGAAGAGCTGTTCGCAACGACCCCTGAGCCAG	1476
Db	480	GluProLeuThrSerLeuLysSerLeuPheGlySerAspProLeuSerGln	496

Search completed: September 1, 2005, 23:50:51
Job time : 399.412 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 20:59:32 ; Search time 266.651 Seconds
(without alignments)
9259.839 Million cell updates/sec

Title: US-09-475-704A-4
Perfect score: 1509
Sequence: 1 atggggcgccgcgcgcacgat.....gcggcccccctgagccagtaa 1509

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1201	79.6	1515	4	US-09-475-515-4 Sequence 4, Appli
2	1201	79.6	4472	4	US-09-475-515-75 Sequence 75, Appl
3	1201	79.6	4608	4	US-09-475-515-76 Sequence 76, Appl
4	1201	79.6	4689	4	US-09-475-515-74 Sequence 74, Appl
5	1201	79.6	4766	4	US-09-475-515-73 Sequence 73, Appl
6	1199	79.5	2031	4	US-09-475-515-7 Sequence 7, Appli
7	1117.8	74.1	1853	4	US-09-475-515-5 Sequence 5, Appli
8	1117.8	74.1	1865	4	US-09-475-515-78 Sequence 78, Appl
9	1117.8	74.1	1865	4	US-09-475-515-79 Sequence 79, Appl
10	1117.8	74.1	4319	4	US-09-475-515-6 Sequence 6, Appli
11	1049.2	69.5	1268	4	US-09-475-515-9 Sequence 9, Appli
12	971.4	64.4	1521	4	US-09-952-060-27 Sequence 27, Appl
13	971.4	64.4	1532	4	US-09-974-702-1 Sequence 1, Appli
14	971.4	64.4	1532	4	US-09-818-443-1 Sequence 1, Appli
15	971.4	64.4	37474	4	US-09-952-060-25 Sequence 25, Appl
16	968.4	64.2	4053	4	US-09-952-060-34 Sequence 34, Appl
17	930.2	61.6	4307	4	US-09-936-572-2 Sequence 2, Appli
18	929.4	61.6	4307	4	US-09-936-572-2 Sequence 2, Appli
19	925.4	61.3	4327	4	US-09-936-572-14 Sequence 14, Appl
20	925.4	61.3	4353	4	US-09-936-572-13 Sequence 13, Appl
21	925.4	61.3	4642	4	US-09-936-572-12 Sequence 12, Appl
22	925.4	61.3	9772	4	US-09-552-950-5 Sequence 5, Appli
23	916.6	60.7	1482	4	US-09-818-443-4 Sequence 4, Appli
24	913.6	60.5	1479	4	US-09-952-060-32 Sequence 32, Appl
25	736.6	48.8	8366	4	US-09-872-733A-6 Sequence 6, Appli
26	734.2	48.7	1496	4	US-09-184-418C-82 Sequence 82, Appl
27	734.2	48.7	8972	4	US-09-184-418C-9 Sequence 9, Appli

28	730.4	48.4	1476	4	US-09-991-258-4	Sequence 4, Appli
29	730.4	48.4	12523	4	US-09-991-258-1	Sequence 1, Appli
30	723.8	48.0	4338	4	US-09-872-733A-1	Sequence 1, Appli
31	712.8	47.2	9913	4	US-09-827-688-11	Sequence 11, Appl
32	709.6	47.0	1485	4	US-09-184-418C-73	Sequence 73, Appl
33	709.6	47.0	9010	4	US-09-184-418C-8	Sequence 8, Appli
34	708	46.9	1476	4	US-09-184-418C-100	Sequence 100, App
35	708	46.9	8959	4	US-09-184-418C-11	Sequence 11, Appl
36	663.6	44.0	1486	4	US-09-184-418C-38	Sequence 38, Appl
37	663.6	44.0	8992	4	US-09-184-418C-4	Sequence 4, Appli
38	651.8	43.2	8968	4	US-09-184-418C-1	Sequence 1, Appli
39	649.4	43.0	1479	4	US-09-184-418C-12	Sequence 12, Appl
40	642.2	42.6	7399	2	US-08-418-848A-9	Sequence 9, Appli
41	642.2	42.6	9709	2	US-08-188-583-5	Sequence 5, Appli
42	642.2	42.6	9709	3	US-08-388-353-1	Sequence 1, Appli
43	642.2	42.6	9709	3	US-08-488-551B-1	Sequence 15, Appl
44	642.2	42.6	9709	3	US-09-309-572-15	Sequence 15, Appl
45	642.2	42.6	9709	4	US-09-718-096-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-09-475-515-4
; Sequence 4, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karlin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475, 515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: HIV-Gag
US-09-475-515-4

Query Match 79.6%; Score 1201; DB 4; Length 1515;
Best Local Similarity 88.5%; Pred. No. 2.8e-173;
Matches 1351; Conservative 0; Mismatches 140; Indels 36; Gaps 3;

QY	1	ATGGGGCGCCCGCCAGCATCCTGCGCGCGAGAACTGAGCAAGTGAGAGATCCGC	60
Db	7	ATGGGGCGCCCGCCAGCATGCTGAGCGCGCGAGAGCTGAGCAAGTGAGAGATCCGC	66
QY	61	CTGGCGCCCGCGCGCAAGAGCACTACATGCTGAAGCACCTGGTGGCGCAGCCGCGAG	120
Db	67	CTGGCGCCCGCGCGCAAGAGTACAAAGCTGAAGCACATGCTGGCGCAGCCGCGAG	126
QY	121	CTGAGGCGCTTCCCTGAACCCCGGCTGCTGAGACCGCGAGGCGTGAAGCAGATC	180
Db	127	CTGAGGCGCTTCCCGTGAACCCCGGCTGCTGAGACCGAGAGGCGTGAAGCAGATC	186
QY	181	ATGAAGCAGCTGAGCCCGCTGAGACCGGAGCAGGAGCTGCGAGCTGTACAAC	240
Db	187	CTGGCCAGCTGAGCCCGCTGAGACCGGAGCAGGAGGAGCTGCGAGCTGTACAAC	246

QY 361 GCCGACGCA-----AGGTGACCAAGACTACCCCATGTCAGAACTG 405
Db 3299 GCCGCCGCAACCGCAACAGCAGCCAGGTGAGCCAGAACTACCCCATGTCAGAACTG 3358
QY 406 CAGGCGCAGATGTCACCAAGCCATCAGCCCCCGCACCCCTGAACGCTGGGTGAAGTG 465
Db 3359 CAGGCGCAGATGTCACCAAGCCATCAGCCCCCGCACCCCTGAACGCTGGGTGAAGTG 3418
QY 466 ATCGAGGAAGGCTTCAGCCCCGAGGTGATCCCATGTTACCGCCCTGAGCGAGGCG 525
Db 3419 GTGAGAGGAAGGCTTCAGCCCCGAGGTGATCCCATGTTACCGCCCTGAGCGAGGCG 3478
QY 526 GCCACCCCCCAGGACCTGAACAGATGTTGAACACCGTGGGGGGCCACAGGCGCCATG 585
Db 3479 GCCACCCCCCAGGACCTGAACAGATGTTGAACACCGTGGGGGGCCACAGGCGCCATG 3538
QY 586 CAGATGCTGAAGGACCACTCAACGAGAGGCGCCGAGTGGGACCGCTGCACCCCGTG 645
Db 3539 CAGATGCTGAAGGAGACCACTCAACGAGAGGCGCCGAGTGGGACCGCTGCACCCCGTG 3598
QY 646 CAGGCGGCGCGCGTGGCCCCCGGCAAGATGCGCGACCCCGCGGCAAGCATCGCCGCG 705
Db 3599 CAGGCGGCGCGCGTGGCCCCCGGCAAGATGCGCGACCCCGCGGCAAGCATCGCCGCG 3658
QY 706 GCCACCAAGACCCCTGCAAGAGCATCGCCTGATGACCAAGCAACCCCGCTGCGCCGTG 765
Db 3659 ACCACCAAGACCCCTGCAAGAGCATCGCCTGATGACCAAGCAACCCCGCTGCGCCGTG 3718
QY 766 GCGGACATCTCAAGCGGTGGATCATCTGGGGCTGAACAAGATCGTGGGATGTACAGC 825
Db 3719 GCGGACATCTCAAGCGGTGGATCATCTGGGGCTGAACAAGATCGTGGGATGTACAGC 3778
QY 826 CCCGTAGCATCTCTGACATCCGCCAGGCCCCCAAGAGCCCTTCCGCGACTACGTGAGC 885
Db 3779 CCCACCAAGCATCTCTGACATCCGCCAGGCCCCCAAGAGCCCTTCCGCGACTACGTGAGC 3838
QY 886 CGCTTCTTCAAGACCCCTGCGCGCGCGGAGGAGCCCAAGGAGCTGAAGAACTGATGAGC 945
Db 3839 CGCTTCTTCAAGACCCCTGCGCGCGCTGAGGAGGAGCCCAAGGAGCTGAAGAACTGATGAGC 3898
QY 946 GAGACCTGCTGTGTCAGAAAGCCCAACCCGACTGCAAGACCATCTGCGGCTCTCGGC 1005
Db 3899 GAGACCTGCTGTGTCAGAAAGCCCAACCCGACTGCAAGACCATCTGAAAGGCTCTCGGC 3958
QY 1006 CCGCGCGCACCCCTGAGAGATGATGACCGCCTGCAAGGCGGTGGCGGCCCCCGCAC 1065
Db 3959 CCGCGCGCACCCCTGAGAGATGATGACCGCCTGCAAGGCGGTGGCGGCCCCCGCAC 4018
QY 1066 AAGGCGCGGTGCTGCGCGAGGCGATGAGCCAGG--CCAACAGCTGAACATCATGATG 1122
Db 4019 AAGGCGCGGTGCTGCGCGAGGCGATGAGCCAGGTGAACCCCGGCAACATCATGATG 4078
QY 1123 CAGAAAGCAACTTCAAGGGCCCCCGGCGCAACGTCAAGTGTCTTCACTGCGCAAGAG 1182
Db 4079 CAGCGGCGCAACTTCCGCAACAGCGGAAGACGTCAAGTGTCTTCACTGCGCAAGAG 4138
QY 1183 GGCACATGCGCAAGACTGCGCGCGCCCCCGCAAGAGGGGTGCTGGAAGTGCAGCAAG 1242
Db 4139 GGCACACCGCGCAAGACTGCGCGCGCCCCCGCAAGAGGGGTGCTGCGCTGCGCGCGC 4198
QY 1243 GAGGCGCACAGATGAAGACTGACCGAGCGCCAGGCAACTTCTGCGGCAAGATCTGG 1302
Db 4199 GAGGCGCACAGATGAAGACTGACCGAGCGCCAGGCAACTTCTGCGGCAAGATCTGG 4258
QY 1303 CCCAGGCAAGAGGCGCGCGCGCAACTTCTGCAAGAACCGCAGCGAGCCCGCGCGCC 1362
Db 4259 CCCAGGCAAGAGGCGCGCGCGCAACTTCTGCAAGAGCG--C 4300
QY 1363 ACCGTGCGCACCGCGCGCGCGCGAGAGCTTCCGCTTGAAGAGACCAACCCCGCGCCC 1422
Db 4301 CCGGAGCGCACCGCGCGCGCGAGAGAGCTTCCGCTTGAAGAGAGAACCAACCCCGC 4360

QY 1423 AAGCAGAGCCCAAGAGACCGCGAGCCCTACCGGAGCCCTGACCGCCCTGCGAGCCTG 1482
Db 4361 AGCCAGAAAGAGAGAGCCCATGCAAGAGAGCTGTACCCCTGACAGCCTGCGAGCCTG 4420
QY 1483 TTCGCGAGCGGCGCCCTGAGCCAGTAA 1509
Db 4421 TTCGCGAAGAGAGCCCAAGAGAGTAA 4447

RESULT 3

US-09-475-515-76
; Sequence 76, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARDOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475, 515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 76
; LENGTH: 4608
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; US-09-475-515-76

Query Match 79.6%; Score 1201; DB 4; Length 4608;
Best Local Similarity 88.5%; Pred. No. 2.8e-173;
Matches 1351; Conservative 0; Mismatches 140; Indels 36; Gaps 3;

QY 1 ATGGGCGCGCGCGCGAGCATCTGCGCGCGCGAGAAAGCTGGAACAAGTGGAGAGATCCGC 60
Db 3075 ATGGGCGCGCGCGCGAGCATCTGCGCGCGCGAGAGCTGGAACAAGTGGAGAGATCCGC 3134
QY 61 CTGCGCGCGCGCGCAAGAGCACTATGCTGAAGCACTGTGTGGCGCGAGCCGCGAG 120
Db 3135 CTGCGCGCGCGCGCAAGAGCACTATGCTGAAGCACTGTGTGGCGCGAGCCGCGAG 3194
QY 121 CTGAGGGCTTGCCTGAACCCCGCGCTGCTGAAGACCGCGAGGGCTGCAAGCATC 180
Db 3195 CTGAGCGCTTGCCTGAACCCCGCGCTGCTGAAGACCGAGGGGCTGCGCCAGATC 3254
QY 181 ATGAAGCAGCTGAGCCCGCGCTGAGACCGGCAACGAGAGAGCTGCGAGCTGTACAAC 240
Db 3255 CTGGGCCAGCTGAGCCCGAGCTGAGACCGGCAACGAGAGAGCTGCGAGCTGTACAAC 3314
QY 241 ACCGTGGCAACCTGTACTGCTGAGACCGCGCATGAGTCCGCGCAACAGAGAGGCC 300
Db 3315 ACCGTGGCAACCTGTACTGCTGAGACCGAGCGCATGAGTCAAGAGACCAAGAGAGGCC 3374
QY 301 CTGACAAGATGAGAGAGAGCAACAAGTCCACAGAAAGACCCAGAGGCCAAGAG 360
Db 3375 CTGAGAAGATGAGAGAGAGCAACAAGTCCAAAGAAAGGCCAAGAGGCCCGCGCC 3434
QY 361 GCCGACGCA-----AGTGAAGCAAGACTACCCCATGTCGAGAACTG 405
Db 3435 GCCGCGGCAACCGGCAACAGAGCAAGTGAAGCAAGACTACCCCATGTCGAGAACTG 3494
QY 406 CAGGCGCAGATGTCACCAAGGCGCATGAGCCCGCACCCCTGAAGCGCTGGGTGAAGTG 465

Db 3495 CAGGGCCAGATGTTGCAACGAGCCATCAGCCCCCGCACCTTGAACGCTGGGTGAAGTG 3554

QY 466 ATCGAGAGAGAGGCTTCAGCCCCCGAGGTGATCCCATGTTCAACCGCTTGAGCGAGGC 525

Db 3555 GTGAGAGAGAGAGGCTTCAGCCCCCGAGGTGATCCCATGTTCAACCGCTTGAGCGAGGC 3614

QY 526 GCCACCCCCCAGAGACTTGAACACGATGTTGAACACCGTGGGGCGGCCACAGGCCGCATG 585

Db 3615 GCCACCCCCCAGAGACTTGAACACGATGTTGAACACCGTGGGGCGGCCACAGGCCGCATG 3674

QY 586 CAGATGTTGAAGSACACCATCAACGAGAGGGCCCGAGTGGGACCGCTGCACCCCGTG 645

Db 3675 CAGATGTTGAAGSACACCATCAACGAGAGGGCCCGAGTGGGACCGCTGCACCCCGTG 3734

QY 646 CAGGGCGGGCCCGTGGGCCCCCGGCCAGATGGCGGACCCCCCGCGCAGACATCGCCGGC 705

Db 3735 CAGGGCGGGCCCGTGGGCCCCCGGCCAGATGGCGGACCCCCCGCGCAGACATCGCCGGC 3794

QY 706 GCCACCAAGCACCCTGCAAGAGAGATGGCTTGATGACCAAGCAACCCCGCTGGCCCGTG 765

Db 3795 ACCACCAAGCACCCTGCAAGAGAGATGGCTTGATGACCAACCAACCCCGCTGGCCCGTG 3854

QY 766 GGGGACATCTCAAGCGGTGGATCATCTGGGGCTTGAACAAGATCGTGGGATGTACAGC 825

Db 3855 GGGGACATCTCAAGCGGTGGATCATCTGGGGCTTGAACAAGATCGTGGGATGTACAGC 3914

QY 826 CCGGTGACATCTCTGACATCCGCCAGGGCCCCCAAGAGCCCTTCCCGCACTACGTGAGC 885

Db 3915 CCCACCAAGCATCTTGACATCCGCCAGGGCCCCCAAGAGCCCTTCCCGCACTACGTGAGC 3974

QY 886 CGCTTCTTCAAGACCCCTGCGCGCGCGCGAGGAGGCCACCAAGAGCGTGAAGAACTGGATGACC 945

Db 3975 CGCTTCTTCAAGACCCCTGCGCGCGCTGAGGAGGCCACCAAGAGCGTGAAGAACTGGATGACC 4034

QY 946 GAGACCTCTGTTGTGCAAGAACGCCAACCCCGACTGCAAGAACCATCTGCGCGCTTCTGCGC 1005

Db 4035 GAGACCTCTGTTGTGCAAGAACGCCAACCCCGACTGCAAGAACCATCTGCGCGCTTCTGCGC 4094

QY 1006 CCGCGCGCCACCCCTGAGAGAGATGATGACCGCTGCAAGGGCGGTGGCGGGCCCCCGGCAC 1065

Db 4095 CCGCGCGCCACCCCTGAGAGAGATGATGACCGCTGCAAGGGCGGTGGCGGGCCCCCGGCAC 4154

QY 1066 AAGGCGCGCGTGTGGCGGAGGCGATGAGCCAGG--CCAACAGCGTGAACATCATGATG 1122

Db 4155 AAGGCGCGCGTGTGGCGGAGGCGATGAGCCAGGTGACGAACCCCGCGACCATCATGATG 4214

QY 1123 CAGAAGAGCAACTTCAAGGGGCCCCCGCGCAACGTCAAGTCTTCAACTGCGGCAAGAG 1182

Db 4215 CAGCGCGGCAACTTCCGCAACAGCGGAAAGACCGTCAAGTCTTCAACTGCGGCAAGAG 4274

QY 1183 GGGCACATCGCCCAAGAACTGCGCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAG 1242

Db 4275 GGGCACATCGCCCAAGAACTGCGCGCGCCCCCGCAAGAGGGCTGCTGCGCGCTGCGCGC 4334

QY 1243 GAGGGCCACAGATGAAGGAAGTGCACCGAGCGCCAGGCCAACTTCTGGGCAAGATCTGG 1302

Db 4335 GAGGGCCACAGATGAAGGAAGTGCACCGAGCGCCAGGCCAACTTCTGGGCAAGATCTGG 4394

QY 1303 CCGAGCCCAAGAGGCGCGCCCCCGGCAACTTCTGCAGAAACGAGCGAGCCCGCGCGCC 1362

Db 4395 CCGAGCCCAAGAGGCGCGCCCCCGGCAACTTCTGCAGAGCG-----C 4436

QY 1363 ACCGTGCCCCACCGCCCCCGCGCGAGAGCTTCCGCTTGCAGAGAACACCCCGCGCCCC 1422

Db 4437 CCGGAGCCCCACCGCCCCCGAGAGAGCTTCCGCTTGGCGAGAGAAAGACCAACCCCGC 4496

QY 1423 AAGCAGAGCCCAAGGACCGGAGCCCTAACCGGAGCCCTGACCGCGCTGCGAGCCTG 1482

Db 4497 AGCCAGAGAGAGAGCCCATGACAAAGAGCTGTACCCCTGACAGCCTTGGCAGCCTG 4556

QY 1483 TTGCGCAGCGGCCCCCTGAGCCAGTAA 1509

Db 4557 TTGCGCAACGAGCCCGCAGCAGCAGTAA 4583

RESULT 4

US-09-475-515-74

; Sequence 74, Application US/09475515A

; Patent No. 6602705

; GENERAL INFORMATION:

; APPLICANT: BARNETT, Susan

; APPLICANT: ZUR MEGEDE, Jan

; APPLICANT: SRIVASTAVA, Indresh

; APPLICANT: LIAN, Ying

; APPLICANT: HARTOG, Karin

; APPLICANT: LIU, Hong

; APPLICANT: GREER, Catherine

; APPLICANT: SELBY, Mark

; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION

; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES

; FILE REFERENCE: 1621.002

; CURRENT APPLICATION NUMBER: US/09/475, 515A

; CURRENT FILING DATE: 1999-12-30

; NUMBER OF SEQ ID NOS: 90

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 74

; LENGTH: 4689

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:

; OTHER INFORMATION: gp160.modSF162.gag.modSF2

US-09-475-515-74

Query Match 79.6%; Score 1201; DB 4; Length 4689;

Best Local Similarity 88.5%; Pred. No. 2.8e-173;

Matches 1351; Conservative 0; Mismatches 140; Indels 36; Gaps 3;

QY 1 ATGGGCGCGCGCGCCAGCATCTGCGCGCGCGGAGAACTGAGCAAGTGGAGAGATCCGC 60

Db 3156 ATGGGCGCGCGCGCCAGCGTGTGAGCGCGCGCGAGCTGAGCAAGTGGAGAGATCCGC 3215

QY 61 CTGCGCGCGCGCGGCAAGAGCACTACATGCTGAAGCACTGTGTGGGCCAGCGCGAG 120

Db 3216 CTGCGCGCGCGCGGCAAGAGAGTACAAAGCTGAAGCACTGTGTGGGCCAGCGCGAG 3275

QY 121 CTGGAGGGCTTGGCCCTGAACCCCGCGCTGCTGGAAGACCGCGAGGGCTGCAAGCAGATC 180

Db 3276 CTGGAGCGCTTGGCCCTGAACCCCGCGCTGCTGGAAGACCGCGAGGGCTGCGCAGATC 3335

QY 181 ATGAAGCAGCTGACGCGCGCGCTGCAAGCCGCGACCGAGAGCTGCGACGCTGTACAA 240

Db 3336 CTGGGCGCAGCTGACGCGCGCTGCAAGCCGCGACCGAGAGCTGCGACGCTGTACAA 3395

QY 241 ACCGTGGCCACCTGTACTGTGCTGACGCGCGCATGAGGTCCGCGACCAAGAGAGCC 300

Db 3396 ACCGTGGCCACCTGTACTGTGCTGACGCGCATGAGGTCCGCGACCAAGAGAGCC 3455

QY 301 CTGGAACAAGATGAGAGAGAGAGCAACAAGTCCAGCAGAAACCCAGAGCCCAAGAG 360

Db 3456 CTGGAAGAAGATGAGAGAGAGAGCAACAAGTCCAAAGAAAGGCCAGAGCGCGCGCC 3515

QY 361 GCCAGCGCA-----AGGTGAGCCAGAACTAACCCCATCGTGAAGAACTGTG 405

Db 3516 GCCGCGGCAACCGGCAACAGACAGCCAGGTGAGCCAGAACTAACCCCATCGTGAAGAACTGTG 3575

QY 406 CAGGGCCAGATGCTGACACGAGCCATCAGCCCGGCAACCTGAAGCGCTGGTGAAGGTG 465

Db 3576 CAGGGCCAGATGCTGACACGAGCCATCAGCCCGGCAACCTGAAGCGCTGGTGAAGGTG 3635

QY 466 ATCGAGAGAGAGGCTTCAAGCCCGGAGGTGATCCCATGTTCAACCGCGCTGAGAGAGGC 525

Db 3636 GTGAGAGAGAGAGGCTTCAAGCCCGGAGGTGATCCCATGTTCAAGCGCGCTGAGAGAGGC 3695

QY 526 GCCACCCCGCAGACCTGAAGACGATGTTGAACACCGTGGCGGCGCACCAAGCGCGCATG 585

```
Db 3696 GCCACCCCCCAGACCTGAACAGATGTTGAACACCGTGGGGCGGCCACAGGCCGCATG 3755
Qy 586 CAGATGCTGAAGACACCATCAACGAGAGGGCCGCCGAGTGGACCGCTGCACCCCGTG 645
Db 3756 CAGATGCTGAAGAGACCATCAACGAGAGGGCCGCCGAGTGGGACCGCGTGCAACCCGCTG 3815
Qy 646 CAGCGCGCCCGGTGGCCCCCGGCGCAAGTGCAGACCCCGCGGACGACATCGCCGC 705
Db 3816 CAGCGCGCCCGCATCGCCCCCGGCGCAAGTGCAGACCCCGCGGACGACATCGCCGC 3875
Qy 706 GCCACGACACCCCTGACGAGACGATCGCTGATGACGACACCCCGCTGCGCGTG 765
Db 3876 ACCACGACACCCCTGACGAGACGATCGGCTGATGACCAACAACCCCGCATCCCGTG 3935
Qy 766 GCGGACATCTACAAGCGGTGATCATCTGGGCTGAACAAGATCGTGGATGTACAGC 825
Db 3936 GCGGAGATCTACAAGCGGTGATCATCTGGGCTGAACAAGATCGTGGATGTACAGC 3995
Qy 826 CCGGTGACATCTGACATCCGCGGCGCCCAAGAGCCCTTCCGCGACTACGTGAC 885
Db 3996 CCCACGACATCTGACATCCGCGGCGCCCAAGAGCCCTTCCGCGACTACGTGAC 4055
Qy 886 CGCTTCTTCAAGACCTGCGCGCGCGGACGAGGCCACCCAGACGTGAAGACTGATGACC 945
Db 4056 CGCTTCTTCAAGACCTGCGCGCGCGTGAAGAGGCCACGACGTGAAGACTGATGACC 4115
Qy 946 GAGACCTGCTGTGTGACAGACGCCAACCCGACTGCAAGACCATCTGCGCGCTTGCC 1005
Db 4116 GAGACCTGCTGTGTGACAGACGCCAACCCGACTGCAAGACCATCTGGAAGGCTTGCC 4175
Qy 1006 CCGCGCGCCACCTGTGAGAGATGATGACCGCTGCGAGGGCGTGGGGCGCCCGCGCAC 1065
Db 4176 CCGCGCGCCACCTGTGAGAGATGATGACCGCTGCGAGGGCGTGGGGCGCCCGCGCAC 4235
Qy 1066 AAGCGCGCGTGTGCGCGGAGGCGGATGAGCCAGG--CCAACAGCGTGAACATCATGATG 1122
Db 4236 AAGCGCGCGTGTGCGCGGAGGCGGATGAGCCAGGTCGAACCCCGGACCATCATGATG 4295
Qy 1123 CAGAAGACCACTTCAAGGGGCCCCCGGCGCAACGTCAAGTGTCTTCAACTGCGGCAAGAG 1182
Db 4296 CAGCGCGGCAACTTCCGCAACCAAGCGGAAGACCGTCAAGTGTCTTCAACTGCGGCAAGAG 4355
Qy 1183 GGGCACAATGCCCAAGAACTGCGCGCGCCCCCGCAAGAGGGCTGTGGAAGTGGCGCAAG 1242
Db 4356 GGGCACAATGCCCAAGAACTGCGCGCGCCCCCGCAAGAGGGCTGTGCGCGTGGCGCGC 4415
Qy 1243 GAGGGCCACCATGATGAAGGAACTGACCGAGCGCCAGGCCAATTCTGGGCAAGATCTGG 1302
Db 4416 GAGGGCCACCATGATGAAGGAACTGACCGAGCGCCAGGCCAATTCTGGGCAAGATCTGG 4475
Qy 1303 CCCAGCCACAAGGGCGCGCGCGCAACTTCTGCAAGAACCGCAGCGAGCCCGCGCGCC 1362
Db 4476 CCCAGCTACAAGGGCGCGCGCGCAACTTCTGCAAGAGCCG-----C 4517
Qy 1363 ACCGTGCCCCACCGCCCCCGCGCGAGAGCTTCCGCTTGAGAGAGACACCCCGCGCCCC 1422
Db 4518 CCGGAGCCCAACCGCCCCCGGAGGAGCTTCCGCTTGCGCGAGAGAGAACCAACCCCGC 4577
Qy 1423 AAGCAGGAGCCCAAGGACCGCGAGCGCTTACCGGAGCCCTGACCGCCCTGCGAGCTG 1482
Db 4578 AGCCAGAGAGAGAGAGCCATGACAGAGAGCTGTACCCCTGACCAAGCTGCGCAGCTG 4637
Qy 1483 TTGGCAGCGGCGCCCTGAGCCAGTAA 1509
Db 4638 TTGGCAACGACCCCAAGCAGCCAGTAA 4664
```

RESULT 5
US-09-475-515-73
; Sequence 73, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:

```
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MESEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 73
; LENGTH: 4766
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; US-09-475-515-73
```

Query Match 79.6%; Score 1201; DB 4; Length 4766;
Best Local Similarity 88.5%; Pred. No. 2.8e-173;
Matches 1351; Conservative 0; Mismatches 140; Indels 36; Gaps 3;

```
Qy 1 ATGGCGCCCGCGCCAGCATCTGCGCGCGGAGAACTGACAAAGTGGAAGATCCGC 60
Db 3233 ATGGCGCCCGCGCCAGCATCTGCGCGCGGAGAACTGGAAGTGGAAGATCCGC 3292
Qy 61 CTGCGCCCGCGCGCAAGAGCACTACATGCTGAAGCACCTGTGTGGCCAGCCGAG 120
Db 3293 CTGCGCCCGCGCGCAAGAGCACTACATGCTGAAGCACCTGTGTGGCCAGCCGAG 3352
Qy 121 CTGAGGGCTTGCCTGAACTCCCGCGCTGCTGAGAACCGCGAGGGCTGCAAGCATC 180
Db 3353 CTGAGGGCTTGCCTGAACTCCCGCGCTGCTGAGAACCGCGAGGGCTGCGCCAGATC 3412
Qy 181 ATGAAGCAGCTGAGCCCGCGCTGAGAACCGGCAACGAGAGCTGCGAGCTGTACAAC 240
Db 3413 CTGGGCGCAGCTGAGCCCGCGCTGAGAACCGGCAACGAGAGCTGCGAGCTGTACAAC 3472
Qy 241 ACCGTGGCCACCTGTACTGCTGTCAGCGCGCATGAGGTCCGACACCAAGAGGCC 300
Db 3473 ACCGTGGCCACCTGTACTGCTGTCAGCGCGCATGAGGTCCGACACCAAGAGGCC 3532
Qy 301 CTGACAAGATGAGAGAGAGAGCAAGTCCAGCAGAGAACCCAGCAGGCCAAGAG 360
Db 3533 CTGAGAAGATGAGAGAGAGAGCAAGTCCAGCAGAGAACCCAGCAGGCCAAGAG 3592
Qy 361 GCCGACGGCA-----AGGTGAGCCAGAACTACCCCATCTGTGAGAACCTG 405
Db 3593 GCCGCGCGCACCGGCAACAGACGACGAGTGAGCCAGAACTACCCCATCTGTGAGAACCTG 3652
Qy 406 CAGGGCCAGATGTGACACGAGCCATCAGCCCGCGACCTGAACGCTGGTGAAGTG 465
Db 3653 CAGGGCCAGATGTGACACGAGCCATCAGCCCGCGACCTGAACGCTGGTGAAGTG 3712
Qy 466 ATCGAGAGAGAGGCTTCAAGCCCGGAGGTGATCCCATATTTCAACCGCCCTGAGCGAGGC 525
Db 3713 GTGAGAGAGAGAGGCTTCAAGCCCGGAGGTGATCCCATATTTCAAGCGCCCTGAGCGAGGC 3772
Qy 526 GCCACCCCGCAGACCTGAACAGATGTTGAACACCGTGGCGGCCACGAGCCGCATG 585
Db 3773 GCCACCCCGCAGACCTGAACAGATGTTGAACACCGTGGCGGCCACGAGCCGCATG 3832
Qy 586 CAGATGCTGAAGACACCATCAACGAGAGGGCCGCGAGTGGAGCCGCTGCACCCCGTG 645
Db 3833 CAGATGCTGAAGAGACCATCAACGAGAGGGCCGCGAGTGGAGCCGCTGCACCCCGTG 3892
```



```

QY 766 GCGACATCTACAAGCGGTGATCATCTGGGCTGAACAAGATCGTGGATGTACAGC 825
    |||
Db 787 GCGAGATCTACAAGCGGTGATCATCTGGGCTGAACAAGATCGTGGATGTACAGC 846
QY 826 CCCGTGACATCTGACATCCGCCAGAGGCCCAAGAGCCCTCCGCACTACGTGAGC 885
    |||
Db 847 CCCACAGCATCTGACATCCGCCAGAGGCCCAAGAGCCCTCCGCACTACGTGAGC 906
QY 886 CGCTTCTTCAAGACCTGCGCGCCGAGAGGCCACCCAGACGTGAAGAAGTGAATGACC 945
    |||
Db 907 CGCTTCTTCAAGACCTGCGCGCTGAGAGGCCAGACGTGAAGAAGTGAATGACC 966
QY 946 GAGACCTGCTGTGACAGACGCCAACCCGACTGCAAGACCATCTCGCGCTCTCGGC 1005
    |||
Db 967 GAGACCTGCTGTGACAGACGCCAACCCGACTGCAAGACCATCTCGCGCTCTCGGC 1026
QY 1006 CCCGCGCCACCTTGAGAGAGATGATGACCCCTGCGAGGGCGTGGCGGCCCGGCCAC 1065
    |||
Db 1027 CCCGCGCCACCTTGAGAGAGATGATGACCCCTGCGAGGGCGTGGCGGCCCGGCCAC 1086
QY 1066 AAGGCCCGCTGTGCGCCAGGCGGATGAGCCAGG---CCAACAGCGTGAACATCATGATG 1122
    |||
Db 1087 AAGGCCCGCTGTGCGCCAGGCGGATGAGCCAGGAGTGAACACCCGCGACATCATGATG 1146
QY 1123 CAGAAGACCACTTCAAGGGGCCCCCGGCGCAACGTCAAGTCTTCAACTGCGGCAAGAG 1182
    |||
Db 1147 CAGCGCGGCAACTTCCGCAACCAAGCGGAAGACGTCAAGTCTTCAACTGCGGCAAGAG 1206
QY 1183 GGGCACTGCGCAAGAACTGCGCGCGCGCGCGCGCAAGAAAGGGCTGCTGGAAGTGGCGCAAG 1242
    |||
Db 1207 GGGCACTGCGCAAGAACTGCGCGCGCGCGCGCGCAAGAAAGGGCTGCTGGAAGTGGCGCGC 1266
QY 1243 GAGGCGCCACGATGAAGAACTGCAACCGAGCGCCAGGCGCACTTCTGGGCAAGATCTGG 1302
    |||
Db 1267 GAGGCGCCACGATGAAGAACTGCAACCGAGCGCCAGGCGCACTTCTGGGCAAGATCTGG 1326
QY 1303 CCCAGCCCAAGAGGCGCGCGCGCGCACTTCTGCAAGAACCGCAGCGCGCGCGCGCGCC 1362
    |||
Db 1327 CCCAGCTCAAGAGGCGCGCGCGCGCACTTCTGCAAGAGCGG-----C 1368
QY 1363 ACCGTGCGCCACCGCGCGCGCGCGCGAGACTTCCGCTTGAAGAGACCAACCCCGCGCGCC 1422
    |||
Db 1369 CCCGAGCCCAAGAGGCGCGCGCGCGAGAGACTTCCGCTTGGGCGAGAGAGACCAACCCCGC 1428
QY 1423 AAGCAGAGAGCCCAAGAGAGCGGAGCGCTTACCGCGAGCGCGCTGACCGCGCGAGCGCTG 1482
    |||
Db 1429 AGCCAGAGAGAGAGAGCGGAGCGCTTACCGCGAGCGCGCTGACCGCGCGAGCGCTG 1488
QY 1483 TTGGGAGCGGCGCGCGCGCGAGT 1507
    |||
Db 1489 TTGGGAGAGAGAGCGCGCGCGAGT 1513

```

RESULT 7

```

US-09-475-515-5
; Sequence 5, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30

```

```

; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1853
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-475-515-5

Query Match      74.1%; Score 1117.8; DB 4; Length 1853;
Best Local Similarity 85.1%; Pred. No. 1e-160;
Matches 1299; Conservative 0; Mismatches 192; Indels 36; Gaps 3;

QY 1 ATGGGCGCCCGCGCCAGCATCTGCGCGCGCGGAGAACTGGAACAAGTGGGAGAAAGATCCGC 60
    |||
Db 7 ATGGGCGCCCGCGCCAGCGGTGCTGAGCGCGCGCGAGCTGGAACAAGTGGGAGAAAGATCCGC 66
QY 61 CTGCGCGCCCGCGCGGCAAGAGCACTACATGCTGGAAGCACTGTGTGTGGCCAGCGCGGAG 120
    |||
Db 67 CTGCGCGCCCGCGCGGCAAGAGCACTACATGCTGGAAGCACTGTGTGTGGCCAGCGCGGAG 126
QY 121 CTGAGAGGCTTGGCTTGAACCCCGGCTGTGAGAACCGCGAGGGCTGCAAGCAGATC 180
    |||
Db 127 CTGAGAGGCTTGGCTTGAACCCCGGCTGTGAGAACCGCGAGGGCTGCGCGCAGATC 186
QY 181 ATGAAGCAGCTGACAGCCCGCTGAGAACCGGCAACCGAGAGCTGCGCAGCTGTACAAC 240
    |||
Db 187 CTGGGCGCAGCTGACAGCCCGCTGAGAACCGGCAACCGAGAGAGCTGCGCAGCTGTACAAC 246
QY 241 ACCGTGGCCACCTGTGTACTGCGTGAACCGCGGCACTGAGGTCCGCGCAACCAAGAGGCC 300
    |||
Db 247 ACCGTGGCCACCTGTGTACTGCGTGAACCGCGGCACTGAGGTCCGCGCAACCAAGAGGCC 306
QY 301 CTGGAACAAGTGAAGAGAGAGCAGAACAAAGTCCAGCAGAAAGACCCAGCGCCAAAGAG 360
    |||
Db 307 CTGGAACAAGTGAAGAGAGAGCAGAACAAAGTCCAGCAGAAAGAGGCCCAAGCGCGCGCC 366
QY 361 GCCGACGGCA-----AGTGAAGCCAGAACTAACCCCATCTGTGACAGAACTG 405
    |||
Db 367 GCCGCGGCAACCGGCAACGAGCAGGTGAGCCAGAACTAACCCCATCTGTGACAGAACTG 426
QY 406 CAGGCGCAGATGTGCAACGAGGCCATCAGCCCGCGCAACCTGAACGCTGGTGAAGGTG 465
    |||
Db 427 CAGGCGCAGATGTGCAACGAGGCCATCAGCCCGCGCAACCTGAACGCTGGTGAAGGTG 486
QY 466 ATCGAGAGAGAGGCTTCAAGCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCAGAGGC 525
    |||
Db 487 GTGAGAGAGAGAGGCTTCAAGCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCAGAGGC 546
QY 526 GCCACCCCCCAGAGACTGAACAGATGTTGAACAACCGTGAGGCGCCACAGGCGGCATG 585
    |||
Db 547 GCCACCCCCCAGAGACTGAACAGATGTTGAACAACCGTGAGGCGCCACAGGCGGCATG 606
QY 586 CAGATGCTGAAGAGACCATCAACGAGAGAGGCGCGAGTGGAGCCGCTGACCCCGTG 645
    |||
Db 607 CAGATGCTGAAGAGACCATCAACGAGAGAGGCGCGAGTGGAGCCGCTGACCCCGTG 666
QY 646 CAGGCGCGCCCGCTGCGCGCGCGCGCAGATGCGGCAACCCCGCGCAGCAGCATGCGCGGC 705
    |||
Db 667 CAGGCGCGCCCGCTGCGCGCGCGCGCAGATGCGGCAACCCCGCGCAGCAGCATGCGCGGC 726
QY 706 GCCACGAGCAGCAGCTGAGAGAGAGATGCGCTGATGAACAGCAACCCCGTGCGCGTG 765
    |||
Db 727 ACCACGAGCAGCAGCTGAGAGAGAGATGCGCTGATGAACCAACCCCGCATCCCGTG 786
QY 766 GCGCATCTACAAGCGGTGATCATCTGGGCTGAACAAGATCGTGGATGTACAGC 825
    |||
Db 787 GCGGATCTACAAGCGGTGATCATCTGGGCTGAACAAGATCGTGGATGTACAGC 846
QY 826 CCGTGAGCATCTGAGACATCCGCCAGGCGCCCAAGAGAGCGCTTCCGCGACTACGTGAGC 885
    |||

```



```
Db      847  CCCACGACATCCTGACATCCGCCAGGGCCCCAAGAGCCCTTCCGGCACTACGTGAC  906
QY      886  CGCTTCTTAAGACCCCTGCGCGCCGAGAGGCCACCAGACGTGAAGACTGATGACC  945
Db      907  CGCTTCAACAAGACCCCTGCGCGCTGAGAGAGCCAGCCAGACGTGAAGAACTGATGACC  966
QY      946  GAGACCTGCTGTGTGAGAACGCCAACCCGACTGCAAGAACCATCTCGCGCTCTCGGC  1005
Db      967  GAGACCTGCTGTGTGAGAACGCCAACCCGACTGCAAGAACCATCTGAGAGCTCTCGGC  1026
QY      1006  CCGCGCGCCACCCCTGAGAGATGATGACCGCTGCCAGGGCGTGGGGCCCCCGGCCAC  1065
Db      1027  CCGCGCGCCACCCCTGAGAGATGATGACCGCTGCCAGGGCGTGGGGCCCCCGGCCAC  1086
QY      1066  AAGCGCGCGCTGTGCGCGAGCGATGAGCCAGG---CCAACAGCGTGAACATCATGATG  1122
Db      1087  AAGCGCGCGCTGTGCGCGAGCGATGAGCCAGGTGACGAACCCGGCCACCATCATGATG  1146
QY      1123  CAGAAAGCAACTTCAAGGGCCCCCGGCGCAACGTCAAGTGTCTCACTGCGGCAAGAG  1182
Db      1147  CAGCGCGGCAACTTCCGCAACCAAGCGGAAGACCGTCAAGTGTCTCACTGCGGCAAGAG  1206
QY      1183  GGGCAATGCGCCAAAGACTGCGCGCGCGCGCCGCAAGAAAGGCTGTGGAAGTGGCGGCAAG  1242
Db      1207  GGGCAACCGCCAGGAAGTCCGCGCGCGCGCGCCGCAAGAAAGGCTGTGCGCGCGCGC  1266
QY      1243  GAGGGCCACCAAGATGAAGAACTGCAACCGAGCGCCAGGCCAACTTCTGGGCAAGATG  1302
Db      1267  GAAGGACCAACAATGAAGAAATTGCACTGAGAGACAGGCTAATTTTAAAGGAAGATG  1326
QY      1303  CCCAGCCACAAGGGCGCGCGCGCACTTCTGCAAGAACCGCAGCGAGCCCGCGCGCCC  1362
Db      1327  CTTCTTAACAAGGGAAGGCGCAGGGAATTTCTTCAAGACAG-----A  1368
QY      1363  ACCGTGCCACCGCGCGCGCGCGCGAGCTTCCGTTTGAAGAGACCAACCCCGCGCCC  1422
Db      1369  CCAGAGCCACAAGCGCGCCCAAGAAAGAGCTTCAAGTTTGGGAGGAGAAAACAACTTCC  1428
QY      1423  AAGCAGAGGCCCAAGAGACCGCGAGCCCTAACCGGAGCCCTGACCGCGCTGCGACCTG  1482
Db      1429  TCTCAGAAGCAGAGCGCGATAGACAAGAACTGTATCTTTAATCTTCCCTCAGATCACTC  1488
QY      1483  TTCGGCAGCGGCGCGCGCTGAGCCAGTAA  1509
Db      1489  TTTGGCAACGACCCCTCTGTACAGTAA  1515
```

```
RESULT 8
US-09-475-515-78
; Sequence 78, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475, 515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 78
; LENGTH: 1865
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
; OTHER INFORMATION: Description of Artificial Sequence: GPI
US-09-475-515-78
```

```
Query Match      74.1%; Score 1117.8; DB 4; Length 1865;
Best Local Similarity 85.1%; Pred. No. 1e-160;
Matches 1299; Conservative 0; Mismatches 192; Indels 36; Gaps 3;
```

```
QY      1  ATGGGCGCGCGCCAGCATCTGCGCGCGGAGAACTGGAAGTGGAGAAATCCGC  60
Db      13  ATGGGCGCGCGCCAGCGTGTGAGCGGCGGAGAGCTGGAAGTGGAGAAATCCGC  72
QY      61  CTGCGCGCGCGCGCAAGAACTACATGCTGAAGCACTGTGTGGGCCAGCGCGAG  120
Db      73  CTGCGCGCGCGCGCAAGAACTACATGCTGAAGCACTGTGTGGGCCAGCGCGAG  132
QY      121  CTGAGGGCTTGCCCTGAACCCCGCGCTGTGAGAACCGCGAGGGCTGCAAGCATC  180
Db      133  CTGAGCGCTTGCCCTGAACCCCGCGCTGTGAGAACCAAGAGGGCTGCGCCAGATC  192
QY      181  ATGAAGCAGCTGAGCCCGCGCTGAGAACCGGCAACGAGAGCTGCGAGCTGTAAAC  240
Db      193  CTGGGCGAGCTGAGCCCAAGCTGAGAACCGGCAACGAGAGCTGCGAGCTGTAAAC  252
QY      241  ACCGTGGCAACCTGTACTGCTGTGACAGCGCGCATGAGGTCCGAGACCAAGAGGCC  300
Db      253  ACCGTGGCAACCTGTACTGCTGTGACAGCGCGCATGAGGTCAAGAGACCAAGAGGCC  312
QY      301  CTGGAACAAGATGAGAGAGAGAGAAACAATGCCAGCAGAAACCGAGCGCCAAAGAG  360
Db      313  CTGGAAGAAGATGAGAGAGAGAGAAACAATGCCAAAGAGGCCAGCAGCGCGCC  372
QY      361  GCCGACGCA-----AGGTAGCCAGAACTAACCCCATCTGTGAGAACTGTG  405
Db      373  GCCGCGCGCACCGGCAACAGACAGACAGTGAGCCAGAACTACCCCATCTGTGAGAACTGTG  432
QY      406  CAGGGCCAGATGTGACACCAAGCCATCAGCCCGCAACCTGTAACGCTGGTGAAGGTG  465
Db      433  CAGGGCCAGATGTGACACCAAGCCATCAGCCCGCAACCTGTAACGCTGGTGAAGGTG  492
QY      466  ATCGAGAGAAAGCCTTCAAGCCCGAGGTGATCCCATGTTCAACCGCCCTGAGAGGGC  525
Db      493  GTGAGAGAAAGCCTTCAAGCCCGAGGTGATCCCATGTTCAAGCGCCCTGAGAGGGC  552
QY      526  GCCACCCCGCAGACCTGAACAGATGTTGAACACGTTGGCGGCCACCAAGCGCCATG  585
Db      553  GCCACCCCGCAGACCTGAACAGATGTTGAACACGTTGGCGGCCACCAAGCGCCATG  612
QY      586  CAGATGCTGAAGACACCATCAACGAGAGCGCGCGAGTGGGACCGCTGACCCCGTG  645
Db      613  CAGATGCTGAAGAGAACCATCAACGAGAGCGCGCGAGTGGGACCGCTGACCCCGTG  672
QY      646  CAGCGCGCGCGCGCGCGCGCGCGAGATGCGGGAACCCCGCGGCAAGCATCGCGGC  705
Db      673  CAGCGCGCGCGCGCGCGCGCGCGAGATGCGGGAACCCCGCGGCAAGCATCGCGGC  732
QY      706  GCCACCAAGACCCCTGCAAGAGAGATGCGCTGATGACCAAGCAACCCCGCTGCCGTG  765
Db      733  ACCACCAAGACCCCTGCAAGAGAGATGCGCTGATGACCAAGCAACCCCGCTGCCGTG  792
QY      766  GGGCAGATCTAAGCGGTGATCATCTGCGCTGAACAAGATCGTGGATGTACAGC  825
Db      793  GGGCAGATCTAAGCGGTGATCATCTGCGCTGAACAAGATCGTGGATGTACAGC  852
QY      826  CCGGTGAGCATCTGGAATCGCGCGAGGCGCCCAAGAGAGCCCTTCCGCACTACGTGAC  885
Db      853  CCCACCAAGCATCTGGAATCGCGCGAGGCGCCCAAGAGAGCCCTTCCGCACTACGTGAC  912
QY      886  CGCTTCTTAAGACCCCTGCGCGCGAGAGGCCAACCAAGAGAGTGAAGAACTGATGACC  945
Db      913  CGCTTCTTAAGACCCCTGCGCGCTGAGAGGCCCAAGAGAGTGAAGAACTGATGACC  972
QY      946  GAGACCTGTGTGTGAGAACGCCAACCCGACTGCAAGAACATCTGCGCGCTCTCGGC  1005
```


Db 973 GAGACCTGCTGTGTCAGAAAGCCAAACCCGACTGCAAGACCATCTGAAAGGCTCTGGCC 1032
 QY 1006 CCGGCGCCACCTTGAGAGATGATACCGCTGCCAGGGCGTGGGCGGCCCGGCCAC 1065
 Db 1033 CCGGCGCCACCTTGAGAGATGATACCGCTGCCAGGGCGTGGGCGGCCCGGCCAC 1092
 QY 1066 AAGGCGCGCTGTGGCGGAGCGGATGAGCCAGG--CCAACAGCGTGAACATCATGATG 1122
 Db 1093 AAGGCGCGCTGTGGCGGAGCGGATGAGCCAGGTGAACCCGCGCAACATCATGATG 1152
 QY 1123 CAGAAAGCAACTTCAAGGGCCCCCGCGCAACGTCAAGTCTTCAACTGCGCAAGAG 1182
 Db 1153 CAGCGCGCAACTTCCGCAACAGCGGAAGACCGTCAAGTCTTCAACTGCGCAAGAG 1212
 QY 1183 GGGCACATCGCCAAAGACTGCGCGCGCCCCCGCAAGAGGGCTGTGGAAGTGGCGCAAG 1242
 Db 1213 GGGCACACCGCCAGGAAGTCCGCGCGCCCCCGCAAGAGGGCTGTGCGCGCGC 1272
 QY 1243 GAGGGCCACAGATGAAGAGCTGCAACGAGCGCCAGCCAACTTCTGGGCAAGATCTGG 1302
 Db 1273 GAAGGACCAAAATGAAGATTGACTGAGAGACAGGCTAATTTTAAAGGAAGATCTGG 1332
 QY 1303 CCGAGCCACAAGGGCGCGCCCCCGCAACTTCTGAGAAACCGAGCGAGCCCGCGCCCC 1362
 Db 1333 CTTCTTAACAAGGGAAGGCCAGGGAATTTTCTTCAAGAGAG-----A 1374
 QY 1363 ACCGTGCCACCGCGCCCCCGCGAGAGCTTCCGTTGAGAGACCAACCCCGCGCCC 1422
 Db 1375 CCAGAGCCAACAGCCCAACAGAGAGCTTCAAGTTTGGAGAGAAAAACAACTCCC 1434
 QY 1423 AAGCAGAGCCCAAGAGACCGGAGCCCTAACCGGAGCCCTGACCGCCTGCGAGCCTG 1482
 Db 1435 TCTCAGAAGCAGAGCGGATAGACAAGAACTGTATCTTAACTTCCCTCAGATCACTC 1494
 QY 1483 TTGCGCAGCGGCCCCCTGAGCCAGTAA 1509
 Db 1495 TTTGGCAACGACCCCTCGTCAACAGTAA 1521

RESULT 9
 US-09-475-515-79
 ; Sequence 79, Application US/09475515A
 ; Patent No. 6602705
 ; GENERAL INFORMATION:
 ; APPLICANT: BARNETT, Susan
 ; APPLICANT: ZUR MEGEDE, Jan
 ; APPLICANT: SRIVASTAVA, Indresh
 ; APPLICANT: LIAN, Ying
 ; APPLICANT: HARTOG, Karin
 ; APPLICANT: LIU, Hong
 ; APPLICANT: GREER, Catherine
 ; APPLICANT: SELBY, Mark
 ; APPLICANT: WALKER, Christopher
 ; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
 ; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
 ; FILE REFERENCE: 1621.002
 ; CURRENT APPLICATION NUMBER: US/09/475,515A
 ; NUMBER OF SEQ ID NOS: 90
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 79
 ; LENGTH: 1865
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: GP2
 US-09-475-515-79

Query Match 74.1%; Score 1117.8; DB 4; Length 1865;
 Best Local Similarity 85.1%; Pred. No. 1e-160;
 Matches 1299; Conservative 0; Mismatches 192; Indels 36; Gaps 3;

QY 1 ATGGCGCGCGCGCCAGCATCTGCGCGCGGAGAACTGGACAAGTGGGAGAAATCCGC 60

Db 13 ATGGCGCGCGCGCGCCAGCGTGTGAGCGCGCGGAGAGCTGGAACAAGTGGAGAAATCCGC 72
 QY 61 CTGCGCCCCCGCGCAAGAGCACTACATGCTGAAGCACTGTGTGGGCCAGCCCGCAG 120
 Db 73 CTGCGCCCCCGCGCAAGAGCACTACATGCTGAAGCACTGTGTGGGCCAGCCCGCAG 132
 QY 121 CTGAAGGCTTCCGCTTGAACCCCGCGCTGTGGAAGACCCCGAGGGCTGCAAGCAATC 180
 Db 133 CTGAAGCGCTTCCGCTTGAACCCCGCGCTGTGGAAGACCAAGCAGGGCTGCCGCAATC 192
 QY 181 ATGAAGCAGCTGAGAGCCCGCTGCAAGACCGGCAACGAGAGCTGCGCAGCTGTACAAC 240
 Db 193 CTGGGCCAGCTGAGAGCCCAAGCTTGCAGAGCCGCGAGCGAGAGCTGCGCAGCTGTACAAC 252
 QY 241 ACCGTGGCCACCTGTACTGCGTGCACGCGCGCATGAGGTCCGCGACACCAAGAGGCC 300
 Db 253 ACCGTGGCCACCTGTACTGCGTGCACAGCGCATTCAGCTCAAGGAACCAAGAGAGGCC 312
 QY 301 CTGGAACAAGATCGAGAGAGAGCAACAAGTCCCAAGCAAGAACCCAGAGGCCAAGAG 360
 Db 313 CTGGAAGAAGATCGAGAGAGAGCAACAAGTCCCAAGGAAGAGGCCAAGAGAGGCC 372
 QY 361 GCGGACGGCA-----AGTGAGCCAGAACTACCCCATCTGTGCAAACTTG 405
 Db 373 GCGCGCGCAACCGGCAACAGAGCCAGGTGAGCCAGAACTACCCCATCTGTGCAAACTTG 432
 QY 406 CAGGGCCAGATGTGTGCAACAGAGCATCAGCCCCCGCACCTGGAACGCTGGGTGAAGTG 465
 Db 433 CAGGGCCAGATGTGTGCAACAGAGCATCAGCCCCCGCACCTGGAACGCTGGGTGAAGTG 492
 QY 466 ATCGAGAGAAAGGCTTCAAGCCCCGAGGTGATCCCATGTTCAACGCGCTGAGCGAGGGC 525
 Db 493 GTGAGAGAAAGGCTTCAAGCCCCGAGGTGATCCCATGTTCAACGCGCTGAGCGAGGGC 552
 QY 526 GCCACCCCCCAGAGCCTGAACAGATGTTGAACACCGTGGGCGCCACAGGCGCCATG 585
 Db 553 GCCACCCCCCAGAGCCTGAACAGATGTTGAACACCGTGGGCGCCACAGGCGCCATG 612
 QY 586 CAGATGCTGAAGACACCATCAAGAGAGCGCGCGAGTGGGACCGCCTGCAACCCCGTG 645
 Db 613 CAGATGCTGAAGAGACCATCAAGAGAGCGCGCGAGTGGGACCGCCTGCAACCCCGTG 672
 QY 646 CAGGCGCGCGCGTGGCCCCCGGACAGATCGCGGACCCCGCGGACGCAATCGCGCGC 705
 Db 673 CAGGCGCGCGCGTGGCCCCCGGACAGATCGCGGACCCCGCGGACGCAATCGCGCGC 732
 QY 706 GCCACAGACACCTGCAGAGAGAGATCGCTGATGACCAAGCAACCCCGCGTCCCGTG 765
 Db 733 ACCACAGACACCTGCAGAGAGAGATCGCTGATGACCAAGCAACCCCGCATCCCGTG 792
 QY 766 GGCGACATCTAAGCGGTGATCATCTGGGCTTGAACAAGATCGTGCAGATGTACAGC 825
 Db 793 GGCGAGATCTAAGCGGTGATCATCTGGGCTTGAACAAGATCGTGCAGATGTACAGC 852
 QY 826 CCGGTAGCATCTGAGACATCGGCGGAGGCGCCCAAGAGAGCCCTCCGCACTACGTGAC 885
 Db 853 CCGACAGCATCTGAGACATCGGCGGAGGCGCCCAAGAGAGCCCTCCGCACTACGTGAC 912
 QY 886 CGCTTCTTCAAGACCTGCGGCGGAGCAGGCGCAACCCAGAGCTGAAGAACTGATGACC 945
 Db 913 CGCTTCTTCAAGACCTGCGGCGGAGCAGGCGCAACCCAGAGCTGAAGAACTGATGACC 972
 QY 946 GAGACCTGCTGTGCAAGAGCGCAACCCGACTGCAAGACCATCTGCGGCTCTCGGC 1005
 Db 973 GAGACCTGCTGTGCAAGAGCGCAACCCGACTGCAAGACCATCTGAAGGCTCTCGGC 1032
 QY 1006 CCGGCGCCACCTTGAGAGATGATGACCGCTGCCAGGGCGTGGCGGCCCGGCCAC 1065
 Db 1033 CCGGCGCCACCTTGAGAGATGATGACCGCTGCCAGGGCGTGGCGGCCCGGCCAC 1092
 QY 1066 AAGGCGCGCTGTGGCGGAGCGGATGAGCCAGG--CCAACAGCGTGAACATCATGATG 1122

Db 1093 AAGCCCCGCTGCTGCGCCGAGCGCATGAGCCAGGTGACGAACCCGCGCATCATGATG 1152
QY 1123 CAGAGAGCACTTCAAGGGGCCCCCGCGCAACGTCAAGTGTCTTCAACTGCGCAAGAG 1182
Db 1153 CAGCGCGGCACTTCCGCAACCAAGCGGAGAACCGTCAAGTGTCTTCAACTGCGCAAGAG 1212
QY 1183 GGCACATCGCCCAAGAACTGCGCGCGCGCGCGCAAGAGGGGTGTGGAAGTGCAGAG 1242
Db 1213 GGCACACCGCGCAAGAACTGCGCGCGCGCGCGCAAGAGGGGTGTGCGCGCGCGC 1272
QY 1243 GAGGCCACCAAGATGAAGACTGACCCGAGCGCGCAACTTCTGCGCAAGATCTGG 1302
Db 1273 GAAGGACACCAATGAAAGATTGCACTGAGACAGGCTAATTTTAAAGGAAGATCTGG 1332
QY 1303 CCAAGCCACAAGGCGCGCGCGCAACTTCTGCAAGAACCGAGCGAGCGCGCGCGCGC 1362
Db 1333 CTTCTCTACAAGGGAAGCGGCAATTTTCTTCAAGACAG-----A 1374
QY 1363 ACCGTGCCCCACCGCGCGCGCGCGCGCAAGCTTCCGCTTCAAGAGACCAACCGCGCGCGC 1422
Db 1375 CCAAGGCCAACAAGCGCGCGCAACCAAGAGAGCTTCAAGTTTGGGAGAGAAACAACCTCC 1434
QY 1423 AAGCAGAGAGCCCAAGGACCGGAGCGCTTACCGGAGCGCGCTGACCGCGCTGCGAGCTG 1482
Db 1435 TCTCAGAGAGCAGAGAGCGGATAGACAGAGAACTGTATCTTTAACTTCCCTCAGATCACTC 1494
QY 1483 TTCGGCAGCGCGCGCGCGCTGAGCCAGTAA 1509
Db 1495 TTGGGCAACGAGCCCTCTGTACAGTAA 1521

RESULT 10

US-09-475-515-6
; Sequence 6, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 4319
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: HIV-gag-polymerase
US-09-475-515-6

Query Match 74.1%; Score 1117.8; DB 4; Length 4319;
Best Local Similarity 85.1%; Pred. No. 1e-160;
Matches 1299; Conservative 0; Mismatches 192; Indels 36; Gaps 3;
QY 1 ATGGGCGCGCGCGCGCGCATCTGCGCGCGCGAGAGCTGGACAAAGTGGAGAAAGATCCGC 60
Db 7 ATGGGCGCGCGCGCGCGCGCTGTGAGCGCGCGCGAGCTGGACAAAGTGGAGAAAGATCCGC 66
QY 61 CTGCGCGCGCGCGCGCGCAAGAGCACTGCTGAAGCACTGTGTGTGGCCAGCGCGGAG 120
Db 67 CTGCGCGCGCGCGCGCGCAAGAGCAAGTACAAAGCTGAAGCACTGTGTGTGGCCAGCGCGGAG 126

QY 121 CTGAGGGCTTCCGCTTGAACCCCGCGCTGTCTGAGAGCCCGCGAGGGGTGCAAGCATC 180
Db 127 CTGAGCGCTTCCGCTTGAACCCCGCGCTGTCTGAGAGCCAGGAGGGGTGCGCCAGATC 186
QY 181 ATGAAGCAGTTCAGCCCGCGCTGACAGCCCGCACCGAGAGCTGCGAGCTGTACAAC 240
Db 187 CTGGGCCAGCTGAGCCCGAGCTGACAGCCCGCAGGAGAGCTGCGAGCTGTACAAC 246
QY 241 ACCGTGGCCACCTGTACTGCGTGCACGCGCGCATGAGGTTCGCGACCAAGAGAGGCC 300
Db 247 ACCGTGGCCACCTGTACTGCGTGCACAGCGCATGACGTCACAAGACCAAGAGAGGCC 306
QY 301 CTGACACAAGATGAGAGAGAGAGACAACAAGTCCAGCAGAAAGACCAGAGCCCAAGAG 360
Db 307 CTGAGAGAAGATGAGAGAGAGAGACAACAAGTCCAGAGAAAGGCCACAGAGCGCGCGCC 366
QY 361 GCGGACGGCA-----AGTGAACCAAGAACTAACCCCATCTGTGAGAACTG 405
Db 367 GCGCGCGGCAACCGGCAACAGAGCCAGGTGAGCCAGAACTAACCCCATCTGTGAGAACTG 426
QY 406 CAGGGCCAGATGTGTGCAACCGGCACTCAGCCCGCGCAACCGCTGGAACGCTGGGTGAAGTG 465
Db 427 CAGGGCCAGATGTGTGCAACCGGCACTCAGCCCGCGCAACCGCTGGAACGCTGGGTGAAGTG 486
QY 466 ATCGAGAGAAAGGCTTCAAGCCCGGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGGC 525
Db 487 GTGAGAGAAAGGCTTCAAGCCCGGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGGC 546
QY 526 GCCACCCCGCAGAGCTGAACAAGATGTTGAACAACCGTGGGCGGCCACAGAGCGGCATG 585
Db 547 GCCACCCCGCAGAGCTGAACAAGATGTTGAACAACCGTGGGCGGCCACAGAGCGGCATG 606
QY 586 CAGATGCTGAAGAGACACCATCAACGAGAGGCGCGCGAGTGGGACCGGCTGCAACCCCGTG 645
Db 607 CAGATGCTGAAGAGAGACCATCAACGAGAGGCGCGAGTGGGACCGGCTGCAACCCCGTG 666
QY 646 CAGGCGCGCGCGCGCGCGCGCGCGAGATGCGCGAACCCCGCGCGCAGCGACATCGCGCGC 705
Db 667 CAGGCGCGCGCGCGCGCGCGCGCGAGATGCGCGAACCCCGCGCGCAGCGACATCGCGCGC 726
QY 706 GCCACCGACACCTGTGAGAGAGAGATCGCTTGATGACCAAGAACCCCGCGTGCGCGTG 765
Db 727 ACCACCGACACCTGTGAGAGAGAGATCGCTTGATGACCAAGAACCCCGCGTGCGCGTG 786
QY 766 GCGGACATCTAACAGCGGTGATCATCTGCGCGCTGAACAAGATCGTGGGATGTACAGC 825
Db 787 GCGGAGATCTAACAGCGGTGATCATCTGCGCGCTGAACAAGATCGTGGGATGTACAGC 846
QY 826 CCGGTGAGCATCTGACATCG 885
Db 847 CCGACCGACATCTGACATCG 906
QY 886 CGCTTCTTCAAGACCTGTGCGCGCGCGAGCAGGCGCAACCGAGCGGTGAAGAACTGATGACC 945
Db 907 CGCTTCTTCAAGACCTGTGCGCGCGCGAGCAGGCGCAACCGAGCGGTGAAGAACTGATGACC 966
QY 946 GAGACCTGTGTGTGCAAGAGCGCAACCCCGCACTGCAAGACCATCTGCGCGCTCTGGC 1005
Db 967 GAGACCTGTGTGTGCAAGAGCGCAACCCCGCACTGCAAGACCATCTGAGGCTCTGGC 1026
QY 1006 CCGGCGCGCACCTGTGAGAGAGATGATGACCGCTTCCAGGCGGTGGCGCGCGCGCGCAC 1065
Db 1027 CCGGCGCGCACCTGTGAGAGAGATGATGACCGCTTCCAGGCGGTGGCGCGCGCGCGCAC 1086
QY 1066 AAGCGCGCGGTGCGCGCGAGTGAAGCGAAG--CCAAGCGGTGAACATCATGATG 1122
Db 1087 AAGCGCGCGGTGCGCGCGAGTGAAGCGAAGCGAAGCGCGCGCGCGCGCGCGCGCGCGCG 1146
QY 1123 CAGAAAGCAACTTCAAGGCGCGCGCGCGCAAGTCAAGTGTCTCAACTGCGCGCAAGAG 1182
Db 1147 CAGCGCGGCAACTTCCGAACCAAGCGGAAGACGTCAAGTGTCTCAACTGCGCGCAAGAG 1206
QY 1183 GCGCACATCGCCAAAGAACTGCGCGCGCGCGCGCGCAAGAGGCGTGTGAAGTGGCGCAAG 1242


```
RESULT 12
US-09-952-060-27
; Sequence 27, Application US/09952060
; Patent No. 6733993
; GENERAL INFORMATION:
; APPLICANT: Emin, Emilio A.
; APPLICANT: Youil, Rima
; APPLICANT: Betz, Andrew J.
; APPLICANT: Chen, Ling
; APPLICANT: Kaslow, David C.
; APPLICANT: Silver, John W.
; APPLICANT: Toner, Timothy J.
; APPLICANT: Casimiro, Danilo R.
; TITLE OF INVENTION: ENHANCED FIRST GENERATION ADENOVIRUS
; TITLE OF INVENTION: VACCINES EXPRESSING CODON OPTIMIZED HIV1-GAG, POL, NEF AND
; TITLE OF INVENTION: MODIFICATIONS
; FILE REFERENCE: 20747Y
; CURRENT APPLICATION NUMBER: US/09/952,060
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: PCT/US01/28861
; PRIOR FILING DATE: 2001-09-14
; PRIOR FILING DATE: 2001-09-14
; PRIOR FILING DATE: 2001-09-07
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/233,180
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 1521
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Codon optimized DNA encoding human HIV-1 gag
US-09-952-060-27

Query Match      64.4%; Score 971.4; DB 4; Length 1521;
Best Local Similarity 79.6%; Pred. No. 1.4e-138;
Matches 1210; Conservative 0; Mismatches 281; Indels 30; Gaps 4;
```

```
QY 472 GAGAGGCTTTCAGCCCCGAGGTGATCCCCATGTTACCGCCCTGAGCGAGGGCCACC 531
   |||||
Db 481 GAGAGGCTTTCAGCCCCGAGGTGATCCCCATGTTCTGCGCCCTGTGAGGGTGCACC 540
QY 532 CCCAGGACCTGAACAAGATTTGAACACCGTGGGCGGCCACCAAGGCCCATGAGATG 591
   |||||
Db 541 CCCAGGACCTGAACAACCATGCTGAACACAGTGGGGGGCCATGAGCTGCCATGACAGATG 600
QY 592 CTGAAGGACACCATCAACGAGAGGCGCCGAGTGGGACCCTGCACCCCGTGCAGGCC 651
   |||||
Db 601 CTGAAGGACACCATCAATGAGAGGCTGTGATGGGACAGGCTGCATCTGTGCACGCT 660
QY 652 GGGCCCGTGGCCCCCGCCAGATGCGCGACCCCGCGGACGACATGCGCGCGCCACC 711
   |||||
Db 661 GGGCCCATTTGCCCCCGCCAGATGAGGAGCCAGGGGCTTGACATTTGCTGGCACACC 720
QY 712 AGCACCTGACAGAGCAGATCGCCTGGATGACCAACCCCCCGTGCCTGGGGCGAC 771
   |||||
Db 721 TCACCCCTCAGAGCAGATTTGGCTGGATGACCAACACCCCCCATCTCTGGGGGAA 780
QY 772 ATCTCAAGCGGTGATCATCTGGGCTGAACAAGATGTTGGGATGTACAGCCCCGTG 831
   |||||
Db 781 ATCTCAAGAGGTGATCATCTGGGCTGAACAAGATTTGTGAGGATGTACTCCCCACC 840
QY 832 AGCATCCTGACATCCGCCAGGGCCCCCAAGAGCCCTTCGCGACTACGTGACCGCTTC 891
   |||||
Db 841 TCATTCCTGACATCAGGACGGCCCCCAAGAGCCCTTCAGGACTATGTGACAGTTTC 900
QY 892 TTCAGAGCCCTGGCGCCGAGCGACGGCCACCAAGAGCTGAAGAACTGGATGACGAGACC 951
   |||||
Db 901 TACAAGACCCTGAGGGCTGAGCGAGGCTCCAGAGAGGTGAAGAACTGGATGACAGAGACC 960
QY 952 CTGCTGTGACAGAACGCCAACCCCGACTGCAAGACCATCTGCGGCTCTCGGCCCGGC 1011
   |||||
Db 961 CTGCTGTGACAGATGCAACCTGACTGCAAGACCATCTGAGAGCCCTGGGCCCTGCT 1020
QY 1012 GCCACCTTGAGAGATGATGACCGCTGCGAGGGCGTGGCGGCCCGGCCACAAGGCC 1071
   |||||
Db 1021 GCCACCTTGAGAGATGATGACAGCTGCCAGGGGGTGGGGGCCCTGTCAACAAGGCC 1080
QY 1072 CGCGTGTGGCGGAGGGCGATGAGCCAGG---CCAACAGCTGAACATCATGATGACAGAG 1128
   |||||
Db 1081 AGGGTGTGGCTGAGGGCCATGTTCCAGGTGACCACTCCGCCACCATCATGATGACAGAGG 1140
QY 1129 AGCAACTTCAAGGGCCCCCGGCCAAGCTCAAGTCTTCAACTGCGGCAAGAGGGCCAC 1188
   |||||
Db 1141 GGCAACTTCAAGGAACCAAGAGAAAGACAGTGAAGTCTTCAACTGTGGCAAGGTGGCCAC 1200
QY 1189 ATGCCCAAGAACTGCGCGCGCCCCCGCCAGAAAGAGGCTGTGAAAGTGGGCAAGAGGGC 1248
   |||||
Db 1201 ATTGCCAAGAACTGTAGGGCCCCCAAGAAAGAGGCTGTGAAAGTGTGGCAAGAGGGC 1260
QY 1249 CACCAATGAAGACTGCAACGAGCGCCCAAGCAACTTCTGGGCAAGATCTGGCCCAAGC 1308
   |||||
Db 1261 CACCAATGAAGACTGCAATGAAGAGGCAAGGCCAATTCTTGGGCAAAATCTGGCCCTCC 1320
QY 1309 CACAAGGGCGCGCGGCAACTTCTGCAAGAACCGCAAGAGCGCGCGCGGCCCAAGCGTG 1368
   |||||
Db 1321 CACAAGGGCAGGCTTGCAACTTCTCAGTCCAGGCGCTGAGGCCCAAGGCCCTCCGAG 1380
QY 1369 CCCACCGCGCGCGCGAGAGCTTCCGCTTGAAGAGAGACACCCCGGCCCAAGCAG 1428
   |||||
Db 1381 GAGTCTTTC-----AGTTTGGGAGAGAGAGACACCCCGCCAGAGAGCAG 1428
QY 1429 GAGCCCAAGAGACCGGAGCCCTTACCGGAGCCCTGACCGCCCTGCGAGCCTGTTCGGC 1488
   |||||
Db 1429 GAGCCCATTTGAACAAGAGCTGTAC-----CCCTTGGCTTCCCTGAGGTCTCTGTTGGC 1482
QY 1489 AGCGGCCCCCTGAGCCAGTAA 1509
   |||||
Db 1483 AACGACCCCTCTCTCCAGTAA 1503
```

```

RESULT 13
US-09-974-702-1
; Sequence 1, Application US/09974702
; Patent No. 6696291
; GENERAL INFORMATION:
; APPLICANT: Shiver, John W.
;           Davies, Mary Ellen
;           Freed, Daniel C.
;           Liu, Margaret A.
;           Perry, Helen C.
; TITLE OF INVENTION: Synthetic HIV Gag Genes
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/974,702
; FILING DATE: 09-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/017,981
; FILING DATE: 03-FEB-1998
; APPLICATION NUMBER: US60/037,854
; FILING DATE: 07-FEB-1997
; APPLICATION NUMBER: GB9705040.5
; FILING DATE: 12-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hand, J. Mark
; REGISTRATION NUMBER: 36,545
; REFERENCE/DOCKET NUMBER: 19730
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-3905
; TELEFAX: 732-594-4720
;
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1532 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-974-702-1

Query Match      64.4%; Score 971.4; DB 4; Length 1532;
Best Local Similarity 79.6%; Pred. No. 1.4e-138;
Matches 1210; Conservative 0; Mismatches 281; Indels 30; Gaps 4;

```

```

Db      190 CTGGGCCAGCTCCAGCCCTCCCTGCAACAGGCTGTAGAGAGCTGAGCTCCCTGTACAAC 249
QY      241 ACCGTGGCCACCCTGTACTGCGTGCACGCCGSCATCGAGTCCCGGACACCAAGAGGCC 300
Db      250 ACAGTGGCTACCTCTGTACTGTGTGACACAGAGATGTATGTGAAGACACCAAGAGGCC 309
QY      301 CTGCAAGATCGAGAGAGAGACAGAACAGTCCCAAGCAGAAAGCCAGCAGC----- 353
Db      310 CTGGAAGAAAGATTGAGAGAGAGAGCAGAACAGTCCAGAAAGAGGCCAGCAGCTGTGCT 369
QY      354 --CAAGAGGCCCAAGGTTAGCCAGAACTAACCCCATCTGTGACAGAACTTGCAGGGC 411
Db      370 GGCACAGGCCAATCCAGCCAGGTGTCCAGAACTAACCCATTGTGACAGAACTTCCAGGGC 429
QY      412 CAGATGTGACACGAGGCCATCAGCCCCCGCACCTGAAAGCCTTGGGTGAAGTGTATCGAG 471
Db      430 CAGATGTGACACGAGGCCATCTCCCCCGAGCCTGAATGCTTGGGTGAAGTGTGAG 489
QY      472 GAGAAAGGCTTACGCCCCGAGGTGATCCCATGTTTCAACGGCCTGAGCGAGGGGCCACC 531
Db      490 GAGAAAGGCTTCTCCCTGAGGTGATCCCATGTTTCTTGCCCTGTGAGGGTGCACC 549
QY      532 CCCCAAGACCTTGAACACGATGTTGAACACCGTGGGGCGCCACAGGCCCATGACAGATG 591
Db      550 CCCAGGACCTGAACACCATGCTGAACACAGTGGGGGGCCATCAGGCTGCATGACAGATG 609
QY      592 CTGAAGGACACCATCAACGAGAGAGGCCCGCGAGTGGGACCGCTGCAACCCCGTGCAGGCC 651
Db      610 CTGAAGGAGACCATCAATGAGAGAGGCTGCTGAGTGGGACAGGCTGATCCTGTGACGCT 669
QY      652 GGGCCCGTGGCCCCCGGCGAGATGCGGACCCCGCGGAGCGCATGCGCCGGCCACC 711
Db      670 GGGCCCATGCCCCCGGCGAGATGAGGAGCCAGGGGCTTGAACATGCTGGACACC 729
QY      712 AGCACCTGACAGAGCAGATGCGCTGATGACCAACAACCCCGTGGCCGTGGGCGAC 771
Db      730 TCCACCTCCAGAGCAGATGCTGATGATGACCAACAACCCCGTGGGCGAA 789
QY      772 ATCTCAAGCGGTGATCATCTGGGCTGAACAAGATCGTGGATGTACAGCCCGTG 831
Db      790 ATCTCAAGAGGTGATCATCTGGGCTGAACAAGATGTGAGATGTACTCCCCACC 849
QY      832 AGCATCTGACATCCGCCAGGGCCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTC 891
Db      850 TCCATCTGACATCAGGACAGGCCCCCAAGAGCCCTTACGGACTATGTGACAGTTTC 909
QY      892 TTCAAGACCTGCGCGCGGAGCAGGCCACCCAGAGAGTGAAGAACTGGATGACGAGACC 951
Db      910 TACAAGACCTGAGGGGCTGAGCAGGCTCCCAAGAGAGTGAAGAACTGGATGACGAGACC 969
QY      952 CTGCTGTGACAGACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCGGC 1011
Db      970 CTGCTGTGACAGATGCAACCCCTGACTGCAAGACCATCTGAAGGCCCTGGGCCCTGCT 1029
QY      1012 GCCACCTGAGAGATGATGACCGCCTGCCAGGGCGTGGCGGCCCGGCCACAAGGCC 1071
Db      1030 GCCACCTGAGAGATGATGACAGCCTGCCAGGGGGTGGGGGCCCTGTCAACAAGGCC 1089
QY      1072 CGCGTGTGCGCAGGCGGATGAGCCAGG--CCAACAGCTGAACATCATGATGACAGAG 1128
Db      1090 AGGGTGTGCTGAGGCGCATGATCCAGGTGACCAACTCCGCACCATCATGATGACAGAG 1149
QY      1129 AGCAACTTCAAGGGCCCCCGCGCAAGTCAAGTCTTCAATCGCGCAAGAGGGCCAC 1188
Db      1150 GGCACCTTCAAGAACCAAGAGAGACAGTGAAGTGTCTCACTGTGCAAGTGGGCCAC 1209
QY      1189 ATGCCAAGAACTGCCGCGCCCCCGCAAGAGGGCTGTGAAGTGGCGCAAGAGGGC 1248
Db      1210 ATTGCCAAGAACTGTAGGGCCCCCAAGAGAGGGCTGTGAAGTGTGGCAAGAGGGC 1269
QY      1249 CACCAATGAAGACTGCAACGAGCGCCAGGCCAATTCTTGGGCAAGATCTGGCCAGC 1308

```


Db 1270 CACGAGTGAAGACTGCATGAGAGCAGGCCAACTTCTGGGCAAAATCTGGCCCTCC 1329

QY 1309 CACAAGGGCCGCGCCGCAACTTCTGCGAAGAACCGCAGCGAGCCCGCCGCCCAACCGTG 1368

Db 1330 CACAAGGGCAGGCTGCGCAACTTCTCCAGTCCAGCGCTGAGCCCAAGCCCTCCCGAG 1389

QY 1369 CCAACGCGCCCGCCCGCAGAGCTTCCGCTTGCAGAGAACACCCCGCCCGCCCAAGCAG 1428

Db 1390 GAGTCTCTC-----AGTTTGGGGAGAGAGAACCAACCCCAAGCAGAGCAG 1437

QY 1429 GAGCCCAAGAGACCGGAGCCCTACCGCGAGCCCTGACCGCCCTGGCGACGCTGTTCCGC 1488

Db 1438 GAGCCCATTTGACAAAGAGCTGTAC-----CCCTGGCCTCCCTGAGGTCCCTGTTGGC 1491

QY 1489 AGCGGCCCCCTGAGCCAGTAA 1509

Db 1492 AACGACCCCTCCTCCAGTAA 1512

RESULT 14

US-09-818-443-1

; Sequence 1, Application US/09818443

; Patent No. 6787351

; GENERAL INFORMATION:

; APPLICANT: Chen, Ling

; APPLICANT: Silver, John W.

; APPLICANT: Betz, Andrew J.

; APPLICANT: Casimiro, Danilo R.

; APPLICANT: Caulfield, Michael J.

; APPLICANT: Chastain, Michael A.

; APPLICANT: Emini, Emilio A.

; TITLE OF INVENTION: ADENOVIRUS CARRYING GAG GENE HIV VACCINE

; FILE REFERENCE: 20440YIA

; CURRENT APPLICATION NUMBER: US/09/818,443

; PRIOR FILING DATE: 2001-09-13

; PRIOR APPLICATION NUMBER: PCT/US00/18332

; PRIOR FILING DATE: 2000-07-03

; PRIOR APPLICATION NUMBER: 60/148,981

; PRIOR FILING DATE: 1999-08-13

; PRIOR APPLICATION NUMBER: 60/142,631

; PRIOR FILING DATE: 1999-07-06

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 1532

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Optimized human HIV-1 gag ORF

US-09-818-443-1

Query Match 64.4%; Score 971.4; DB 4; Length 1532;
Best Local Similarity 79.6%; Pred. No. 1.4e-138;
Matches 1210; Conservative 0; Mismatches 281; Indels 30; Gaps 4;

QY 1 ATGGGCGCCCGCGCAGCATCTGCGCGCGGAGAAAGCTGACAAGTGGGAGAAGATCCGC 60

Db 10 ATGGGTCTAGGCTTCTGTCTGTCTGTGTGTGAGCTGACAGTGGGAGAAGATCAGG 69

QY 61 CTGCGCCCCCGCGGCAAGAGCACTACATGTGAAGCACTGTGTGGGCCAGCCGCGAG 120

Db 70 CTGAGGCTGTGTGCAAGAGAGTACAAGTAAAGCATTTGTGTGGCCTCCAGGAG 129

QY 121 CTGAGGGCTTTCGCTGAAACCCCGCTGTGAGACCGCGAGGCTGCAAGCATC 180

Db 130 CTGAGAGGTTTGTGTGAACCTGTGCTGTGAGACCTCTGAGGGGTGAGGCGATC 189

QY 181 ATGAAGAGCTGCAGCCCGCTGCAGACCGGCAAGAGAGTGCAGAGCTGTACAAC 240

Db 190 CTGGGCCAGCTCCAGCCCTCCCTGCACAACAGGCTCTGAGGAGCTGAGGTCCCTGTACAAC 249

QY 241 ACCGTGGCACCTGTACTGTGCTGCAGCCCGGATCGAGTCCGCGACCAAGAGGCC 300

Db 250 ACAGTGGCTACCCCTGTACTGTGTGACACAGAAAGATTGATGTGAAGACACCAAGAGGCC 309

QY 301 CTGGAACAAGATGAGGAGAGAGAGAAACAAGTCCAGCAGAAAGACCAGAGC----- 353

Db 310 CTGGAAGAAAGATTGAGAGAGAGAGAAACAAGTCCAAGAAAGAGGCCAGAGGCTGTGCT 369

QY 354 --CAAGAGGCCGACGGCAAGGTGAGCCAGAACTACCCCATGTGTGAGAACTGCAGGGC 411

Db 370 GGCAACAGCAACTCCAGCCAGGTGTCCAGAACTAACCCATGTGTGAGAACTCCAGGGC 429

QY 412 CAGATGTGCACAGGCCCATCAGCCCCCGCAGACCTTGAAAGCCTGGGTGAAGGTGATCGAG 471

Db 430 CAGATGTGCACAGGCCCATCAGCCCCCGCAGACCTTGAAAGCCTGGGTGAAGGTGATCGAG 489

QY 472 GAGAAGGCTTCAAGCCCGAGGTGATCCCAATGTTCAACGCCCTGAGCGAGGCGCCACC 531

Db 490 GAGAAGGCTTCTCCCTGAGGTGATCCCAATGTTCTGTGCCCTGTGAGGGTGCCACC 549

QY 532 CCCCAGACCTGAACAGATGTTGAACACCGTGGGCGGCCACAGCGCCGATGAGATG 591

Db 550 CCCCAGACCTGAACACCATGCTGAACACAGTGGGGGCCATCAGGCTGCCATGAGATG 609

QY 592 CTGAAGGACACCATCAACGAGAGGCGCGCATGTGGACCGCCTGCACCCCGTCAGGCC 651

Db 610 CTGAAGGAGACCATCAATGAGAGGCTGCTGAGTGGGACAGGCTGCATCCTGTGCACGCT 669

QY 652 GGGCCCCGTGGCCCCCGCCAGATGCGGACCCCCCGCGGACGACATCGCCGCGCCACC 711

Db 670 GGGCCCCATGGCCCCCGCCAGATGAGGAGCCAGGGGCTCTGACATTTGCTGGCACACC 729

QY 712 AGCAACCTGACAGAGCAGATCGCTGATGACCAAGCAACCCCCCGTGCCCGTGGCGGAC 771

Db 730 TCCACCTCCAGAGCAGATTTGGCTGATGACCAACAACCCCCCATCCCTGTGGGGGAA 789

QY 772 ATCTACAAGCGGTGATCATCTCTGGGCGCTGAACAAGATCGTGGGATGTACAGCCCCGTG 831

Db 790 ATCTACAAGAGGTGATCATCTCTGGGCGCTGAACAAGATTTGTGAGATGTACTCCCCACC 849

QY 832 AGCATCTGACATCCGCCAGGCGCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTC 891

Db 850 TCCATCTGACATCAGGACAGGCGCCCAAGAGCCCTTCAAGGACTATGTGACAGGTTT 909

QY 892 TTCAAGACCTTGGCGCGCGGAGAGGCGCCAGAGAGCTGAGAACTGGATGACGAGACC 951

Db 910 TACAAGACCTTGAAGGCGCTGAGCAGGCGCTCCAGAGAGGTGAAGAACTGGATGACAGAGACC 969

QY 952 CTGCTGTGTCAGAAAGCCCAACCCCGACTGCAGAACCATCTCTGGCGGCTCTCGGCCCCGGC 1011

Db 970 CTGCTGTGTCAGAAATGCCAACCTTGACTGCAGAACCATCTCTGAAGGCCCTGGGCGCTGCT 1029

QY 1012 GCCAACCCTGAGAGATGATGACCGGCTGCGAGGGCGTGGGCGGCCCCCGCCACAAGGCC 1071

Db 1030 GCCAACCCTGAGAGATGATGACAGCCTGCGAGGGGGTGGGGGGCCCTGTGTACAAGGCC 1089

QY 1072 CGGCTGTGCGCGAGGCGGATGAGCCAGG--CCAACAGCGTGAACATCATGATGACAGAA 1128

Db 1090 AGGCTGTGCTGAGGCGCATGTCCAGGTGACCAACTCCGCAACCATCATGATGACAGAG 1149

QY 1129 AGCAACTTCAAGGCGCGCGCGCAACGTCAAAGTCTTCAAAGTGGCGCAAGAGGGCCAC 1188

Db 1150 GGCAACTTCAAGAAACAGAGAGAAAGTGAAGTGTCTTCAAGTGTGGCAAGGTGGGCCAC 1209

QY 1189 ATGCCCAAGAACTGCGCGCGCGCGCAAGAAAGGCTGTGGAAGTGGCGCAAGAGGGC 1248

Db 1210 ATTGCCAAGAACTGTAGGGCGCGCGCAAGAAAGGCTGTGGAAGTGTGGCAAGAGGGC 1269

QY 1249 CACGAGTGAAGACTGCACCGAGCGCGCAAGCACTTCTGGGCAAGATCTGGCCAGC 1308

Db 1270 CACGAGTGAAGACTGCATGAGAGGAGGCAAGCACTTCTGTGGCAAAATCTGGCCCTCC 1329

QY 1309 CACAAGGCGCGCGCGCAACTTCTGTGAGAAACGAGCGAGCGCGCGCGCCCAACCGTG 1368

Db 1330 CACAAGGCGAGGCTGGCAACTTCTCAGTCAAGGCGCTGAGGCCACAGCCCTTCCGAG 1389

QY 1369 CCCACCGCCCCCCCCCGAGAGCTTCGCTTCGAGGAGACCACCCCGCCCAAGCAG 1428
 ||| |
 Db 1390 GAGTCTTC-----AGGTTGGGAGAGAGAAGACCACCCCAGCCAGAAAGCAG 1437
 QY 1429 GAGCCCAAGACCGCGAGCCCTACCGCGAGCCCTGACCGCCCTGGCAGCCTGTTGGC 1488
 ||||| |
 Db 1438 GAGCCCATTTGACAAAGAGCTGTAC-----CCCTGGCCCTCCCTGAGGTCCCTGTTGGC 1491
 QY 1489 AGCGGCCCCCTGAGCCAGTAA 1509
 ||| |||||
 Db 1492 AACGACCCCTCTCTCCAGTAA 1512

RESULT 15

```

US-09-952-060-25
; Sequence 25, Application US/09952060
; Patent No. 6733993
; GENERAL INFORMATION:
; APPLICANT: Emini, Emilio A.
; APPLICANT: Youil, Rima
; APPLICANT: Betc, Andrew J.
; APPLICANT: Chen, Ling
; APPLICANT: Kaslow, David C.
; APPLICANT: Shiver, John W.
; APPLICANT: Toner, Timothy J.
; APPLICANT: Casimiro, Danilo R.
; TITLE OF INVENTION: ENHANCED FIRST GENERATION ADENOVIRUS
; TITLE OF INVENTION: VACCINES EXPRESSING CODON OPTIMIZED HIV1-GAG, POL, NEF AND
; FILE REFERENCE: 20747Y
; CURRENT APPLICATION NUMBER: US/09/952,060
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: PCT/US01/28861
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/317,814
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/279,056
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/233,180
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 37474
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA encoding pmkRads HIV-1 gag, coding
US-09-952-060-25

```

Query Match	64.4%	Score 971.4	DB 4	Length 37474
Best Local Similarity	79.6%	Pred. No. 1.3e-138		
Matches 1210	Conservative	0	Mismatches 281	Indels 30
				Gaps 4
Qy	1	ATGGCGCCCGCGCCAGCATCTGCGCGCGCGAAGCTGACAAAGTGGAGAAATCCGC	60	
Db	1272	ATGGGTGCTAGGGCTTCTGTGCTGTCTGTGTGTAGCTGGACAAAGTGGAGAAATCAGG	1331	
Qy	61	CTGCGCCCCGCGCGCAAGACCTACATGTGAAGCACCTGTGTGGCCAGCCGCGAG	120	
Db	1332	CTGAGGCTGTGGCAAGAAAGTACAAAGCTTAAAGCACATGTGTGGGCTCCAGGGAG	1391	
Qy	121	CTGAGGGCTTTCGCCCTCTGAACCCCGGCGCTGTGAGAGCCGCCAGGGCTGCAACAGATC	180	
Db	1392	CTGAGAGAGTTGTGCTGTGAACCCCTGGCCTGTGAGAGACCTTGAAGGGGTGACGGCAGATC	1451	
Qy	181	ATGAAGCAGCTGCAGCCCGCCTGTGACAGACCGGACCGAGGAGCTGCGCAGCCTGTACAAC	240	
Db	1452	CTGGGCGCAGCTCCAGCCCTCCCTGCAAAACAGGCTCTGAGGAGCTGAGGTCCCTGTACAAC	1511	
Qy	241	ACCGTGCGCACCTGTACTGCGTGCACGCGGCGATCGAGGTCCGCGACACCAAGAGGCC	300	

D	b		1512	ACAGTGGCTACCCTGTACTGTGTGTCACCAGAAAGATTGATGTGAAGGACCAAGAAGGCC	1571
O	y		301	CTGGACAAGATCGAGGAGGAGCAGAACAACTCCAGCAGAAAGACCAGCAGGC-----	353
D	b		1572	CTGGAGAAGATTGAGGAGGAGCAGAACAAAGTTCCAAGAAAGAGGCCAGCAGGCTGTGCT	1631
O	y		354	--CAAGGAGGCCGACGGCAAAGGTGAGCCAGAACTACCCCCATCTGTGCAGAACTTGACGGC	411
D	b		1632	GGCACAGGCAACTTCCAGCCAGGTGTCCAGAACTAACCCCATTTGTGCAAGAACTTCCAGGGC	1691
O	y		412	CAGATGTGCACCAAGGCCATCAGCCCCCGCACCTGAACGCTTGGGTGAAGSTGATCGAG	471
D	b		1692	CAGATGTGCACCAAGGCCATCTCCCCCGACCTGTAATGCTTGGGTGAAGSTGTGAG	1751
O	y		472	GAGAAAGGCTTTCAGCCCCGAGGTGATCCCCTATGTTCAACGCTTGAAGGAGGCGCCACC	531
D	b		1752	GAGAAAGGCTTCTCTCCCTGAGGTGATCCCCTATGTTCTTGCCCTGTCTGAAGGTTGCCACC	1811
O	y		532	CCCCAGGACCTGAACACCGATTTGAACACCGTGGGCGGCCACAGGCCGCGCATGCAAGATG	591
D	b		1812	CCCCAGGACCTGAACACCGATCTGAACACAGTGGGGGCCATCAGGCTGCATGCAAGATG	1871
O	y		592	CTGAAGGACACCATCAACCGAGGAGGCCGCGGAGTGGGACCGCCTGCACCCCGTGCAAGCC	651
D	b		1872	CTGAAGGAGACCATCAATGAGGAGGCTGTGAGTGGGACAGGCTGCATCTGTGCACGCT	1931
O	y		652	GCGCCCGTGGCCCCCGGCAGATGCGCGACCCCGCGCAGCGACATCGCCGGCGCCACC	711
D	b		1932	GCGCCCATTTGCCCCCGGCAGATGAGGAGGCCAGGGGCTCTGACATTTGCTGGCACCAACC	1991
O	y		712	AGCACCTTGCAAGGAGCAGATCGCCTGATGACCAAGCAACCCCGTGCCTGGGCGAC	771
D	b		1992	TCCACCTTCCAGGAGCAGATTTGCTGATGACCAACAAACCCCGCATCTGTGGGGGAA	2051
O	y		772	ATCTCAAGCGGTGATCATCTCTGGGCTGAACAAGATCGTGGATGTAACGCCCGTG	831
D	b		2052	ATCTCAAGAGGTGATCATCTCTGGGCTGAACAAGATTTGAGGATGTACTCCCCACC	2111
O	y		832	AGCATCTTGAATCCGCGCAGGGCCCCCAAGAGCCCTTCCGACTACGTGAACCGCTTC	891
D	b		2112	TCCATCTTGAATCATCAGGCGAGGGCCCCAAGAGCCCTTCAAGGACTATGTGAACAGTTTC	2171
O	y		892	TTCAAGACCTTGC CGCGCCGAGCAGGCCACCAAGACGTGAAGAACTGTGACCGAGACC	951
D	b		2172	TACAAAGACCTTGAAGGGCTGAGCAAGCCCTCCAGAGGTGAAGAACTGTGACAGAGACC	2231
O	y		952	CTGCTGTGTGCAAAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTGGCCCCGGC	1011
D	b		2232	CTGCTGTGTGCAAAATGCCAACCTGACTGCAAGACCATCTGGAAGGCCCTTGGGCCCTGCT	2291
O	y		1012	GCCACCTTGAAGAGATGATGACCGCCTTGCAGAGGGCGTGGCGGCCCGGCACAAAGCC	1071
D	b		2292	GCCACCTTGAAGAGATGATGACCGCTTGCAGAGGGGGTGGGGGGCCCTGTGTCAAAAGCC	2351
O	y		1072	GCGGTGTGGCGGAGGCGATGAGCCAGG---CCAACAGCGTGAACATCATGATGCAGAG	1128
D	b		2352	AGGGTGTGGCTGAGGCCATGTCTCCAGGTGAACCAACTCGGCACCATCATGATGCAGAGG	2411
O	y		1129	AGCAACTTCAAGGGCCCCCGCGCAACGTCAAGTGTCTTCAACTGCGGCAAGAGGGCCAC	1188
D	b		2412	GCGCAACTTCAAGAAACAGAGAAAGACAGTGAAGTGTCTCAACTGTGCAAGTGGGCCAC	2471
O	y		1189	ATCGCCAAAGAACTGCGCGCGCCCCCGCAAGAAAGGCTGTGGAAGTGCAGAGAGGC	1248
D	b		2472	ATTGCCAAAGAACTTAGGGCCCCCAAGAAAGGCTGTGGAAGTGTGCAAGAGAGGC	2531
O	y		1249	CACCAAGATGAAGACTGCACGAGCGCAGGCCCACTTCTTGGGCAAGATCTGGCCCCAGC	1308
D	b		2532	CACCAGATGAAGACTGCATAAGAGGCAAGGCCCACTTCTTGGGCAAAATCTGGCCCTTC	2591
O	y		1309	CACAAAGGCGCCCCCGCAACTTCTTGCAAAACGCAAGAGCCCCCGCCCCCAACCTTG	1368
D	b		2592	CACAAAGGCGCGCTTGCAACTTCTCTCAAGTCCAGGCTGAGCCCAAGCCCCCTCCGAG	2651

```
QY      1369 CCCACCGCCCCCCCCCGAGAGCTTCCGCTTGAAGAGACACCCCGCCCCCAAGCAG 1428
          |||
Db      2652 GAGTCTTC-----AGTTTGGGGAGAGAGACACCCCAAGCAGAGCAG 2699
          |||

QY      1429 GAGCCCAAGGACCGGAGCCCTACCGGAGACCCCTGACCGCCCTGCGCAGCCTGTTGGC 1488
          |||||
Db      2700 GAGCCCATTGACAAGAGAGCTGTAC-----CCCTGGCCTCCCTGAGGTCCCTGTGGC 2753
          |||||

QY      1489 AGCGGCCCCCTGAGCCAGTAA 1509
          |||
Db      2754 AACGACCCCTCCTCCCAGTAA 2774
          |||
```

Search completed: September 1, 2005, 22:46:26
Job time : 274.651 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 21:10:07 ; Search time 935,802 Seconds
(without alignments)
10560.254 Million cell updates/sec

Title: US-09-475-704A-4
Perfect score: 1509
Sequence: 1 atcgggcggccgcgcagcat.....gcggcccccctgagccagtaa 1509

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 7338684 seqs, 3274456166 residues

Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
- 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
- 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1509	100.0	1509	10	US-09-899-575-4 Sequence 4, Appli
2	1507.4	99.9	1509	10	US-09-967-464-64 Sequence 64, Appl
3	1504.2	99.7	1509	10	US-09-967-464-68 Sequence 68, Appl
4	1489.8	98.7	1509	10	US-09-899-575-21 Sequence 21, Appl
5	1297.6	86.0	1494	10	US-09-899-575-51 Sequence 51, Appl
6	1294.8	85.8	1491	10	US-09-899-575-99 Sequence 99, Appl
7	1276.8	84.6	1479	10	US-09-967-464-63 Sequence 63, Appl

	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	
	1276.8	1275.2	1275.2	1275.2	1275.2	1275.2	1275.2	1275.2	1275.2	1275.2	1273.6	1273.6	1273.6	1273.6	1273.6	1273.6	1273.6	1273.6	1273.6	1272.6	1272.6	1271.6	1271.2	1270.2	1270.2	1260.8	1201	1201	1201	1201	1201	1199.4	1199.4	1199.4	1199.4	1199.4	1199.4	1199.4	
	84.6	84.5	84.5	84.5	84.5	84.5	84.5	84.5	84.5	84.5	84.4	84.4	84.4	84.4	84.4	84.4	84.4	84.4	84.3	84.3	84.3	84.3	84.2	84.2	84.2	83.6	79.6	79.6	79.6	79.6	79.5	79.5	79.5	79.5	79.5	79.5	79.5	79.5	79.5
	1479	3162	3162	3162	3162	3162	3162	3162	3162	3162	2742	2742	2742	2742	2742	2742	2742	2742	2742	2742	2742	2742	2742	2742	2742	1479	1479	1479	1479	1479	1479	1479	1479	1479	1479	1479	1479	1479	1479
	10	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15
	US-09-899-575-3	US-10-190-435-18	US-10-190-435-16	US-10-190-435-19	US-10-190-435-14	US-10-190-435-35	US-10-190-435-34	US-10-190-435-36	US-10-190-435-38	US-10-190-435-17	US-10-190-305A-13	US-10-190-435-20	US-10-190-435-12	US-10-190-435-11	US-10-190-305A-12	US-10-190-305A-83	US-10-190-305A-82	US-10-190-305A-81	US-09-899-575-20	US-10-387-336-4	US-10-387-336-75	US-10-387-336-76	US-10-387-336-74	US-10-387-336-73	US-10-387-336-73	US-10-387-336-73	US-10-387-336-73	US-10-387-336-73	US-10-387-336-73	US-10-387-336-73	US-10-387-336-73	US-10-387-336-73	US-10-387-336-73	US-10-387-336-73	US-10-387-336-73	US-10-387-336-73	US-10-387-336-73	US-10-387-336-73	US-10-387-336-73
	Sequence 3, Appli	Sequence 18, Appl	Sequence 16, Appl	Sequence 19, Appl	Sequence 35, Appl	Sequence 36, Appl	Sequence 34, Appl	Sequence 38, Appl	Sequence 17, Appl	Sequence 13, Appl	Sequence 20, Appl	Sequence 15, Appl	Sequence 9, Appli	Sequence 10, Appl	Sequence 11, Appl	Sequence 12, Appl	Sequence 59, Appl	Sequence 58, Appl	Sequence 82, Appl	Sequence 67, Appl	Sequence 57, Appl	Sequence 81, Appl	Sequence 20, Appl	Sequence 4, Appli	Sequence 75, Appl	Sequence 76, Appl	Sequence 74, Appl	Sequence 73, Appl	Sequence 18, Appl	Sequence 11, Appl	Sequence 17, Appl	Sequence 15, Appl	Sequence 15, Appl	Sequence 15, Appl	Sequence 15, Appl	Sequence 15, Appl	Sequence 15, Appl	Sequence 15, Appl	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-09-899-575-4
; Sequence 4, Application US/09899575
; Publication No. US20030223961A1
; GENERAL INFORMATION:
; APPLICANT: Zur Megele, Jan
; APPLICANT: Barnett, Susan W.
; APPLICANT: Egnelbrecht, Susan
; APPLICANT: van Rensburg, Estrelita Janse
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; FILE REFERENCE: PP01631.102
; CURRENT APPLICATION NUMBER: US/09/899,575
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 09/475,704
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1509
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag
; OTHER INFORMATION: of HIV strain AF110967
US-09-899-575-4

Query Match 100.0%; Score 1509; DB 10; Length 1509;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1509; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGGCGCCCGCCAGCATCTGCGCGGCGAGAAAGTGGACAAGTGGAGAAAGATCCGC 60
Db 1 ATGGGCGCCCGCCAGCATCTGCGCGGCGAGAAAGTGGACAAGTGGAGAAAGATCCGC 60
QY 61 CTGCGCGCCCGCGCGAGAAAGCACTAATGCTGAAGCACTGTGTGGCCAGCCGCGAG 120
Db 61 CTGCGCGCCCGCGCGAGAAAGCACTAATGCTGAAGCACTGTGTGGCCAGCCGCGAG 120
QY 121 CTGAGGGCTTTCGCTCTGAACCCCGGCTGCTGAAGACCGCCGAGGCTGCAAGCAATC 180
Db 121 CTGAGGGCTTTCGCTCTGAACCCCGGCTGCTGAAGACCGCCGAGGCTGCAAGCAATC 180
QY 181 ATGAAGCAGCTGACGCGCCCTGCAAGCCGACCGAGAGCTGCGCAGCCTGTACAAC 240
Db 181 ATGAAGCAGCTGACGCGCCCTGCAAGCCGACCGAGAGCTGCGCAGCCTGTACAAC 240
QY 241 ACCGTGGCCACCCCTGTACTGCGTGCAAGCCGCGCATCGAGTCCGCGACACCAAGAGGCC 300
Db 241 ACCGTGGCCACCCCTGTACTGCGTGCAAGCCGCGCATCGAGTCCGCGACACCAAGAGGCC 300
QY 301 CTGACAAGATCGAGAGGAGAGAGACAAGTCCACAGAGAAGCCAGCAGGCCAAGAG 360
Db 301 CTGACAAGATCGAGAGGAGAGAGACAAGTCCACAGAGAAGCCAGCAGGCCAAGAG 360
QY 361 GCCGACGGCAAGGTGAGCCAGAACTACCCCATCTGTGAGAACTGTGAGGGCCAGATGGTG 420
Db 361 GCCGACGGCAAGGTGAGCCAGAACTACCCCATCTGTGAGAACTGTGAGGGCCAGATGGTG 420
QY 421 CACCAAGCCATCAGCCCGCCGACCCCTGAAAGCTGGTGAAGTGAAGTGAAGAGGCC 480
Db 421 CACCAAGCCATCAGCCCGCCGACCCCTGAAAGCTGGTGAAGTGAAGTGAAGAGGCC 480
QY 481 TTCAAGCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGCGCCACCCCGCAGGAC 540
Db 481 TTCAAGCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGCGCCACCCCGCAGGAC 540
QY 541 CTGAACACGATGTTGAACACCGTGGGGGCCACACAGCGCCCATGAGATGCTGAAGAG 600
Db 541 CTGAACACGATGTTGAACACCGTGGGGGCCACACAGCGCCCATGAGATGCTGAAGAG 600
QY 601 ACCATCAACGAGAGGGCGCGGAGTGGGACCGCTGCAACCCCGTGCAGGGCGGCCGCTG 660
Db 601 ACCATCAACGAGAGGGCGCGGAGTGGGACCGCTGCAACCCCGTGCAGGGCGGCCGCTG 660
QY 661 GCGCGCGCGCGAGATGGCGGACCCCGCGGCGAGCATCGCGCGCGCACCAAGACCTTG 720
Db 661 GCGCGCGCGCGAGATGGCGGACCCCGCGGCGAGCATCGCGCGCGCACCAAGACCTTG 720
QY 721 CAGAGCAGATCGCTGATGACCAAGAACCCCGCTGCGCGCGCATCTACAAG 780
Db 721 CAGAGCAGATCGCTGATGACCAAGAACCCCGCTGCGCGCGCATCTACAAG 780
QY 781 CGGTGATCATCTGGGCTTGAACAAGATCGTGGGATGTACAGCCCGTGAAGCATCTTG 840
Db 781 CGGTGATCATCTGGGCTTGAACAAGATCGTGGGATGTACAGCCCGTGAAGCATCTTG 840
QY 841 GACATCCGCGCGAGGGCGCCCAAGAGGCTTCCGCGACTACGTGACCGCTTCTCAAGACC 900
Db 841 GACATCCGCGCGAGGGCGCCCAAGAGGCTTCCGCGACTACGTGACCGCTTCTCAAGACC 900
QY 901 CTGCGCGCGCGAGAGGCCACCAAGACGTGAAGACTGGATGACCCAGACCTGCTGTG 960
Db 901 CTGCGCGCGCGAGAGGCCACCAAGACGTGAAGACTGGATGACCCAGACCTGCTGTG 960
QY 961 CAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTGCGCGCGCGCGCACCTTG 1020
Db 961 CAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTGCGCGCGCGCGCACCTTG 1020
QY 1021 GAGGAGATGATGACCGCTGCGAGGGGCTGGGGCGCGCGCGCACAGGGCGCGCTG 1080
Db 1021 GAGGAGATGATGACCGCTGCGAGGGGCTGGGGCGCGCGCGCACAGGGCGCGCTG 1080
```

```
QY 1081 GCCGAGCGGATGAGCCAGGCCAACAGCGTGAACATCATGATGCAAGAGCAACTTCAG 1140
Db 1081 GCCGAGCGGATGAGCCAGGCCAACAGCGTGAACATCATGATGCAAGAGCAACTTCAG 1140
QY 1141 GGGCCCCGGCGCAACGTCAAGTCTTCAACTGCGCGCAAGAGGGCCACATCGCCAAAGAC 1200
Db 1141 GGGCCCCGGCGCAACGTCAAGTCTTCAACTGCGCGCAAGAGGGCCACATCGCCAAAGAC 1200
QY 1201 TGCCGGCGCCCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGAGGGCCACCATGTAAG 1260
Db 1201 TGCCGGCGCCCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGAGGGCCACCATGTAAG 1260
QY 1261 GACTGCACCGAGCGCCAGGCCAATTCTGCGCAAGATCTGGGCCAAGCCACCAAGGGCCGC 1320
Db 1261 GACTGCACCGAGCGCCAGGCCAATTCTGCGCAAGATCTGGGCCAAGCCACCAAGGGCCGC 1320
QY 1321 CCGGCGCAACTTCTCTGCAAGAACCGCAGCGAGCCCGCGCCCGCCACCGTCCACCGCCCC 1380
Db 1321 CCGGCGCAACTTCTCTGCAAGAACCGCAGCGAGCCCGCGCCCGCCACCGTCCACCGCCCC 1380
QY 1381 CCGGCGCAAGCTTCTCGTGAAGAGACCAACCCCGCGCCCGCCCAAGCAGAGCCCAAGAG 1440
Db 1381 CCGGCGCAAGCTTCTCGTGAAGAGACCAACCCCGCGCCCGCCCAAGCAGAGCCCAAGAG 1440
QY 1441 CCGGAGCCCTTACCGCGAGCCCGCTGACCGCGCTGCGCAGCCTGTTGGCAGCGGCCCTG 1500
Db 1441 CCGGAGCCCTTACCGCGAGCCCGCTGACCGCGCTGCGCAGCCTGTTGGCAGCGGCCCTG 1500
QY 1501 AGCCAGTAA 1509
Db 1501 AGCCAGTAA 1509
```

```
RESULT 2
US-09-967-464-64
; Sequence 64, Application US/09967464
; Publication No. US20030138453A1
; GENERAL INFORMATION:
; APPLICANT: O'Hagan, Derek
; APPLICANT: Otten, Gills
; APPLICANT: Donnelly, John J.
; APPLICANT: Polo, John M.
; APPLICANT: Barnett, Susan
; APPLICANT: Singh, Mamohan
; APPLICANT: Uimer, Jeffrey
; APPLICANT: Dubensky, Jr., Thomas W.
; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS
; FILE REFERENCE: P16269.004
; CURRENT APPLICATION NUMBER: US/09/967,464
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/236,105
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/315,905
; PRIOR FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64
; LENGTH: 1509
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-09-967-464-64
```

```
Query Match 99.9%; Score 1507.4; DB 10; Length 1509;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1508; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGCGCCCGCGCAGCATCTGCGCGGCGAGAAAGTGGACAAGTGGAGAAAGATCCGC 60
Db 1 ATGGGCGCCCGCGCAGCATCTGCGCGGCGAGAAAGTGGACAAGTGGAGAAAGATCCGC 60
QY 61 CTGCGCGCGCGCGCAAGAGCACTAATGCTGAAGCACTGTGTGGCCAGCCGCGAG 120
Db 61 CTGCGCGCGCGCGCAAGAGCACTAATGCTGAAGCACTGTGTGGCCAGCCGCGAG 120
```



```
QY 121 CTGAGGCTTCGCTTGAACCCCGGCTGTGAGACCGCCGAGGGCTGCAAGCAGATC 180
Db 121 CTGAGGCTTCGCTTGAACCCCGGCTGTGAGACCGCCGAGGGCTGCAAGCAGATC 180
QY 181 ATGAAGCAGCTGCAGCCCGCTGCAGACCGGCACCGAGAGCTGCAGCCTGTACAA 240
Db 181 ATGAAGCAGCTGCAGCCCGCTGCAGACCGGCACCGAGAGCTGCAGCCTGTACAA 240
QY 241 ACCGTGGCCACCCCTGACTGCTGACGCGCGCATCGAGGTCCGCGACACCAAGAGGCC 300
Db 241 ACCGTGGCCACCCCTGACTGCTGACGCGCGCATCGAGGTCCGCGACACCAAGAGGCC 300
QY 301 CTGACAAGATCGAGGAGAGCAGAACAGTCCACAGAGAACCCGACGAGCCAGAGAG 360
Db 301 CTGACAAGATCGAGGAGAGCAGAACAGTCCACAGAGAACCCGACGAGCCAGAGAG 360
QY 361 GCCGACGGCAAGGTGAGCCAGAACTACCCCATCTGTGAGAACTGAGGGCCAGATGGTG 420
Db 361 GCCGACGGCAAGGTGAGCCAGAACTACCCCATCTGTGAGAACTGAGGGCCAGATGGTG 420
QY 421 CACCAAGCCATCAGCCCCCGCACCCCTGAAAGCTGGAAGGTGATCGAGAGAGAGCC 480
Db 421 CACCAAGCCATCAGCCCCCGCACCCCTGGAAGGTGATCGAGAGAGAGCC 480
QY 481 TTCAAGCCCGAGGTGATCCCCCATGTTCACCGCCCTGAGCGGAGGCCACCCCCAGAG 540
Db 481 TTCAAGCCCGAGGTGATCCCCCATGTTCACCGCCCTGAGCGGAGGCCACCCCCAGAG 540
QY 541 CTGAACACGATGTTGAACACCGTGGGCGGCCACCAAGCCGCCATGCAGATGCTGAAGAC 600
Db 541 CTGAACACGATGTTGAACACCGTGGGCGGCCACCAAGCCGCCATGCAGATGCTGAAGAC 600
QY 601 ACCATCAACGAGAGAGCGCCCGAGTGGGACCGCTGCAACCCCTGCAAGCCCGCCCGTG 660
Db 601 ACCATCAACGAGAGAGCGCCCGAGTGGGACCGCTGCAACCCCTGCAAGCCCGCCCGTG 660
QY 661 GCCCGCCCGCAGATGCGCGCACCCCGCGGCAGCGACATCGCCGCGCACCAAGCAGCCTG 720
Db 661 GCCCGCCCGCAGATGCGCGCACCCCGCGGCAGCGACATCGCCGCGCACCAAGCAGCCTG 720
QY 721 CAGGACGAGATCGCTTGGATGACCAAGAACCCCGCTGCGCGCGACATCTACAAAG 780
Db 721 CAGGACGAGATCGCTTGGATGACCAAGAACCCCGCTGCGCGCGACATCTACAAAG 780
QY 781 CGGTGATCATCTGGGCTGAAACAAGATCGTGGGATGTACAGCCCGTGAGCATCTCG 840
Db 781 CGGTGATCATCTGGGCTGAAACAAGATCGTGGGATGTACAGCCCGTGAGCATCTCG 840
QY 841 GACATCCGCGCAGGCGCCCAAGAGCCCTTCGCGCATACGTGAACCGCTTCTTCAAGACC 900
Db 841 GACATCCGCGCAGGCGCCCAAGAGCCCTTCGCGCATACGTGAACCGCTTCTTCAAGACC 900
QY 901 CTGCGCGCCGAGCAGGCGCACCCAGAGCGTGAAGAACTGGATGACCGAGACCTGTGCTG 960
Db 901 CTGCGCGCCGAGCAGGCGCACCCAGAGCGTGAAGAACTGGATGACCGAGACCTGTGCTG 960
QY 961 CAGAACGCCAACCCCGACTGCAAGACATCTGCGCGCTCTCGGCCCGCGCCACCTCG 1020
Db 961 CAGAACGCCAACCCCGACTGCAAGACATCTGCGCGCTCTCGGCCCGCGCCACCTCG 1020
QY 1021 GAGGAGATGATGACCGCTGCGAGGCGTGGCGCGCCCGGCACAAAGCCCGCTGCTG 1080
Db 1021 GAGGAGATGATGACCGCTGCGAGGCGTGGCGCGCCCGGCACAAAGCCCGCTGCTG 1080
QY 1081 GCCGAGCGATGAGCCAGGCGCAACAGCGTGAACATCATGTGACAGAGAGCAACTTCAAG 1140
Db 1081 GCCGAGCGATGAGCCAGGCGCAACAGCGTGAACATCATGTGACAGAGAGCAACTTCAAG 1140
QY 1141 GGCCCGCGCGCAACGTCAGTGTCTTCACTGCGGCAAGAGGGGCCATCGCAAGAAC 1200
Db 1141 GGCCCGCGCGCAACGTCAGTGTCTTCACTGCGGCAAGAGGGGCCATCGCAAGAAC 1200
```

```
QY 1201 TGCCGCGCCCCCGCAAGAGGCTGTGGAAGTGGCGCAAGAGGGCCACCAAGATGAAG 1260
Db 1201 TGCCGCGCCCCCGCAAGAGGCTGTGGAAGTGGCGCAAGAGGGCCACCAAGATGAAG 1260
QY 1261 GACTGCACCGAGCGCCAGGCCAACTTCTCTGGGCAAGATCTGGCCAGCCACAGAGGGCCG 1320
Db 1261 GACTGCACCGAGCGCCAGGCCAACTTCTCTGGGCAAGATCTGGCCAGCCACAGAGGGCCG 1320
QY 1321 CCGGCAACTTCTCTGCAGAAACCGAGCGAGCCCGCCCGCCCAACCGTGCCACCGCCCC 1380
Db 1321 CCGGCAACTTCTCTGCAGAAACCGAGCGAGCCCGCCCGCCCAACCGTGCCACCGCCCC 1380
QY 1381 CCGCGCAGAGCTTCCGCTTGCAGAGAACACCCCGCCCGCCCAAGAGAGCCCAAGAGAC 1440
Db 1381 CCGCGCAGAGCTTCCGCTTGCAGAGAACACCCCGCCCGCCCAAGAGAGCCCAAGAGAC 1440
QY 1441 CCGGAGCCCTACCGCGAGCCCTGACCGCCCTGCGCAGCCTGTTGGCAGCGGCCCTG 1500
Db 1441 CCGGAGCCCTACCGCGAGCCCTGACCGCCCTGCGCAGCCTGTTGGCAGCGGCCCTG 1500
QY 1501 AGCCAGTAA 1509
Db 1501 AGCCAGTAA 1509
```

RESULT 3

```
US-09-967-464-68
; Sequence 68, Application US/09967464
; Publication No. US20030138453A1
; GENERAL INFORMATION:
; APPLICANT: O'Hagan, Derek
; APPLICANT: Otten, Gillis
; APPLICANT: Donnelly, John J.
; APPLICANT: Polo, John M.
; APPLICANT: Barnett, Susan
; APPLICANT: Singh, Mamohan
; APPLICANT: Ulmer, Jeffrey
; APPLICANT: Dubensky, Jr., Thomas W.
; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS
; FILE REFERENCE: P16269.004
; CURRENT APPLICATION NUMBER: US/09/967,464
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/236,105
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/315,905
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68
; LENGTH: 1509
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-09-967-464-68
```

```
Query Match 99.7%; Score 1504.2; DB 10; Length 1509;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1497; Conservative 12; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGGGCGCGCGCCAGCATCTCTGCGCGCGCGGAGAAAGCTGAGCAAGTGGAGAAATCCGC 60
Db 1 ATGGGCGCGCGCCAGCATCTCTGCGCGCGCGGAGAAAGCTGAGCAAGTGGAGAAATCCGC 60
QY 61 CTGCGCGCGCGGCAAGAGCACTACATGTCTGAAGCACTGTGTGGGCCAGCGCGAG 120
Db 61 CTGCGCGCGCGGCAAGAGCACTACATGTCTGAAGCACTGTGTGGGCCAGCGCGAG 120
QY 121 CTGGAAGGCTTGCCCTGAACCCCGGCTGTGAGAGACCGCGGAGGGCTGCAAGCAGATC 180
Db 121 CTGGAAGGCTTGCCCTGAACCCCGGCTGTGAGAGACCGCGGAGGGCTGCAAGCAGATC 180
QY 181 ATGAAGCAGCTGAGCCCGCGCTGCAAGACCGGAGAGCTGCGAGAGCTGCGAGCCTGTACAA 240
Db 181 ATGAAGCAGCTGAGCCCGCGCTGCAAGACCGGAGAGCTGCGAGAGCTGCGAGCCTGTACAA 240
```

```
QY 241 ACCGTGGCCACCCCTGTACTGCGTGCAAGCCGGCATCGAGGTCCGCGACAACCAAGAGGCC 300
Db 241 ACCGTGGCCACCCCTGTACTGCGTGCAAGCCGGCATCGAGGTSCGCGACAACCAAGAGGCC 300
QY 301 CTGGAACAAGATCGAGAGGAGCAGAACAGTCCCAAGAAAGACCACCAAGGCCAAGAG 360
Db 301 CTGGAACAAGATCGAGAGGAGCAGAACAGTCCCAAGAAAGACCACCAAGGCCAAGAG 360
QY 361 GCCGACGGCAAGGTGAGCCAGAACTAACCCCATCTGTGCAAACTTGACAGAGGCCAGATGGTG 420
Db 361 GCCGACGGCAAGGTGAGCCAGAACTAACCCCATCTGTGCAAACTTGACAGAGGCCAGATGGTG 420
QY 421 CACCAAGCCATCAGCCCCCGCACCTTGAAAGCCTGGGTGAAGGTGATCGAGAGAGAGCC 480
Db 421 CACCAAGCCATCAGCCCCCGCACCTTGAAAGCCTGGGTGAAGGTGATCGAGAGAGAGCC 480
QY 481 TTCAGCCCCGAGGTGATCCCATGTTTCAACCGCCCTGAGCGAGGGCGCCACCCCGCAGGAC 540
Db 481 TTCAGCCCCGAGGTGATCCCATGTTTCAACCGCCCTGAGCGAGGGCGCCACCCCGCAGGAC 540
QY 541 CTGAACAAGATGTTGAACAACCGTGGCGGCCACCAAGCCCGCATGCAAGATGCTGAAGAG 600
Db 541 CTGAACAAGATGTTGAACAACCGTGGCGGCCACCAAGCCCGCATGCAAGATGCTGAAGAG 600
QY 601 ACCATCAACGAGAGGCGCCGAGTGGGACCGCCTGCACCCCGTGACAGCCGCGCCGCGCTG 660
Db 601 ACCATCAACGAGAGGCGCCGAGTGGGACCGCCTGCACCCCGTGACAGCCGCGCCGCGCTG 660
QY 661 GCCCGCGCCGAGATGCGCGACGCCCGCGGACAGCAGATCGCGCGCCACCAAGACCCCTG 720
Db 661 GCCCGCGCCGAGATGCGCGACGCCCGCGGACAGCAGATCGCGCGCCACCAAGACCCCTG 720
QY 721 CAGGAGCAGATCGCTGGATGACCAAGCAACCCCGTGCCCGTGCGGCGACATCTACAAG 780
Db 721 CAGGAGCAGATCGCTGGATGACCAAGCAACCCCGTGCCCGTGCGGCGACATCTACAAG 780
QY 781 CGGTGGATCATCTGGGCTTGAACAAGATCGTGGGATGTACAGCCCGTGAGCATCTCTG 840
Db 781 CGGTGGATCATCTGGGCTTGAACAAGATCGTGGGATGTACAGCCCGTGAGCATCTCTG 840
QY 841 GACATCCCGCAGGCGCCCAAGAGAGCCCTTCCGCACTACGTGACCGCTTCTTCAAGACC 900
Db 841 GACATCCCGCAGGCGCCCAAGAGAGCCCTTCCGCACTACGTGACCGCTTCTTCAAGACC 900
QY 901 CTGCGCGCCGAGCAGGCCACCAAGACGTGAAGACTGGATGACCGAGACCCCTGCTGGTG 960
Db 901 CTGCGCGCCGAGCAGGCCACCAAGACGTGAAGACTGGATGACCGAGACCCCTGCTGGTG 960
QY 961 CAGAACGCCAACCCCGACTGCAAGACCATCTGCGGCTCTCGGCCCCGGCGCCACCCCTG 1020
Db 961 CAGAACGCCAACCCCGACTGCAAGACCATCTGCGGCTCTCGGCCCCGGCGCCACCCCTG 1020
QY 1021 GAGGAGATGATGACCGCTGCGAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCTGCTG 1080
Db 1021 GAGGAGATGATGACCGCTGCGAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCTGCTG 1080
QY 1081 GCCGAGGCGATGAGCCAGGCCACAAGCGTGAACATCATGATGCAAGAGCAACTTCAAG 1140
Db 1081 GCCGAGGCGATGAGCCAGGCCACAAGCGTGAACATCATGATGCAAGAGCAACTTCAAG 1140
QY 1141 GGGCCCCGGCGCAAGCTSAAGTGTCTTCAACTGCGGCAAGAGGGGCCCATCGCCAAAGAAC 1200
Db 1141 GGGCCCCGGCGCAAGCTSAAGTGTCTTCAACTGCGGCAAGAGGGGCCCATCGCCAAAGAAC 1200
QY 1201 TGCCGCGCCCCCGCAAGAGGGCTGTGAAGTGGGCAAGAGGGGCCACCAAGATGAAG 1260
Db 1201 TGCCGCGCCCCCGCAAGAGGGCTGTGAAGTGGGCAAGAGGGGCCACCAAGATGAAG 1260
QY 1261 GACTGCAACCGAGCGCCCAACTTCTTGCGCAAGATCTGGCCCAAGGGCGCC 1320
Db 1261 GACTGCAACCGAGCGCCCAACTTCTTGCGCAAGATCTGGCCCAAGGGCGCC 1320
```

```
QY 1321 CCGCGCAACTTCTCTGCAGAACCCGACGAGCCCGCGCCCGCCCAACCGTGCCCAACGCCCCC 1380
Db 1321 CCGCGCAACTTCTCTGCAGAACCCGACGAGCCCGCGCCCGCCCAACCGTGCCCAACGCCCCC 1380
QY 1381 CCGCGCGAGAGCTTCCGCTTGAAGAGAACCAACCCCGCGCCCGCCCAAGAGAGCCCAAGAG 1440
Db 1381 CCGCGCGAGAGCTTCCGCTTGAAGAGAACCAACCCCGCGCCCGCCCAAGAGAGCCCAAGAG 1440
QY 1441 CGCGAGCCCTACCGCGAGCCCGCTGACCGCCCTGCGCAGCCTGTGCGAGCGGCCCCCTG 1500
Db 1441 CGCGAGCCCTACCGCGAGCCCGCTGACCGCCCTGCGCAGCCTGTGCGAGCGGCCCCCTG 1500
QY 1501 AGCCAGTAA 1509
Db 1501 AGCCAGTAA 1509
```

RESULT 4

```
US-09-899-575-21
; Sequence 21, Application US/09899575
; Publication No. US20030223961A1
; GENERAL INFORMATION:
; APPLICANT: Zur Megele, Jan
; APPLICANT: Barnett, Susan W.
; APPLICANT: Egnelbrecht, Susan
; APPLICANT: van Rensburg, Estrelita Janse
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; FILE REFERENCE: PP01631.102
; CURRENT APPLICATION NUMBER: US/09/899,575
; CURRENT FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 09/475,704
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 1509
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag
; OTHER INFORMATION: coding sequence of HIV strain AP10967
US-09-899-575-21
```

Query Match 98.7%; Score 1489.8; DB 10; Length 1509;
Best Local Similarity 99.2%; Pred. No. 3.4e-313;
Matches 1497; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
QY 1 ATGGGGCCCCCGCCAGCATCTGCGCGCGAGAAAGTGAACAAGTGAGAGATCCGC 60
Db 1 ATGGGGCCCCCGCCAGCATCTGCGCGCGAGAAAGTGAACAAGTGAGAGATCCGC 60
QY 61 CTGCGCCCCCGCGCAAGAGCACTACATGCTGAAGCACTGGTGTGGCCAGCGCGAG 120
Db 61 CTGCGCCCCCGCGCAAGAGCACTACATGCTGAAGCACTGGTGTGGCCAGCGCGAG 120
QY 121 CTGAGGGCTTGCCTGAAACCCCGGCTGCTGGAAGACCGCGAGGGCTGCAAGCATC 180
Db 121 CTGAGGGCTTGCCTGAAACCCCGGCTGCTGGAAGACCGCGAGGGCTGCAAGCATC 180
QY 181 ATGAAGCAGCTGCAAGCCCGCTGCAAGCCGGACCGAGAGCTGCGAGCCTGTACAAC 240
Db 181 ATGAAGCAGCTGCAAGCCCGCTGCAAGCCGGACCGAGAGCTGCGAGCCTGTACAAC 240
QY 241 ACCGTGGCCACCCCTGACTGCGTGCAAGCGCGCATCGAGTCCGCGACCAAGAGGCC 300
Db 241 ACCGTGGCCACCCCTGACTGCGTGCAAGCGCGCATCGAGTCCGCGACCAAGAGGCC 300
QY 301 CTGGAACAAGATGAGAGGAGCAGAACAGTCCCAAGAGAGACCCAGAGGCCAAGAG 360
Db 301 CTGGAACAAGATGAGAGGAGCAGAACAGTCCCAAGAGAGACCCAGAGGCCAAGAG 360
QY 361 GCCGACGGCAAGGTGAGCCAGAACTAACCCCATCTGTGCAAGACTTGACAGGGCCAGATG 420
```

```
Db      361 GCGGACGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACTTGAGGGCCAGATGGTG 420
QY      421 CACCAAGGCATCAGCCCCCGCACCCCTGAAAGCTGGGTGAGGTGATCGAGAGAGGCC 480
Db      421 CACCAAGGCATCAGCCCCCGCACCCCTGAAAGCTGGGTGAGGTGATCGAGAGAGGCC 480
QY      481 TTCAGCCCCGAGGTGATCCCATGTTCAACCGCTGAGCGAGGGCGCCACCCCGCAGGAC 540
Db      481 TTCAGCCCCGAGGTGATCCCATGTTCAACCGCTGAGCGAGGGCGCCACCCCGCAGGAC 540
QY      541 CTGAACACGATGTTGAACACCGTGGGGGCCACCAAGCCCGCATGAGATGCTGAAGAC 600
Db      541 CTGAACACCATGCTGAACACCGTGGGGGCCACCAAGCCCGCATGAGATGCTGAAGAC 600
QY      601 ACCATCAACGAGAGGCGCGCGAGTGGGACCGCTGCACCCCGTGACGGCCGCGCGTG 660
Db      601 ACCATCAACGAGAGGCGCGCGAGTGGGACCGCTGCACCCCGTGACGGCCGCGCGTG 660
QY      661 GCGCGCGCGCGAGATGCGCGGACCCCGCGGCGAGCATCGCGCGCGCCACCAAGCCTG 720
Db      661 GCGCGCGCGCGAGATGCGCGGACCCCGCGGCGAGCATCGCGCGCGCCACCAAGCCTG 720
QY      721 CAGGACGAGATCGCTGGATGACCAACCAACCCCGCTGCGCGCGGCGACATCTACAAG 780
Db      721 CAGGACGAGATCGCTGGATGACCAACCAACCCCGCTGCGCGCGGCGACATCTACAAG 780
QY      781 CGGTGATCATCTCTGGGCTTGAACAGATCGTGGGATGTACAGCCCGTGAGCATCTG 840
Db      781 CGGTGATCATCTCTGGGCTTGAACAGATCGTGGGATGTACAGCCCGTGAGCATCTG 840
QY      841 GACATCCGCGCAGGCGCCCAAGAGGCGCTTCCGCGACTACGTGAGCGCTTCTTCAAGAC 900
Db      841 GACATCCGCGCAGGCGCCCAAGAGGCGCTTCCGCGACTACGTGAGCGCTTCTTCAAGAC 900
QY      901 CTGCGCGCGCGAGGAGCCCAACGAGACTGGAAGACTGATGACCGAGACCTGCTGGTG 960
Db      901 CTGCGCGCGCGAGGAGCCCAACGAGACTGGAAGACTGATGACCGAGACCTGCTGGTG 960
QY      961 CAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGCGCGCGCGCGCACCTG 1020
Db      961 CAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGCGCGCGCGCGCACCTG 1020
QY      1021 GAGGAGATGATGACCGCTGCGAGGCGTGGGCGCGCGCGCGCACAGGCGCGCGTGCTG 1080
Db      1021 GAGGAGATGATGACCGCTGCGAGGCGTGGGCGCGCGCGCGCACAGGCGCGCGTGCTG 1080
QY      1081 GCGGAGGCGATGAGCCAGGCCCAACAGCGTGAACTATGATGACAGAGAGCAACTTCAAG 1140
Db      1081 GCGGAGGCGATGAGCCAGGCCCAACAGCGTGAACTATGATGACAGAGAGCAACTTCAAG 1140
QY      1141 GCGCGCGCGCGCAACGTCAAGTGTCTTCACTGCGGCAAGAGGCGCCATCGCCCAAGAAC 1200
Db      1141 GCGCGCGCGCGCAACGTCAAGTGTCTTCACTGCGGCAAGAGGCGCCATCGCCCAAGAAC 1200
QY      1201 TGCAGCGCGCGCGCAAGAGGCGTGTGGAAGTGGCGCAAGGAGGCCACCAAGTGAAG 1260
Db      1201 TGCAGCGCGCGCGCAAGAGGCGTGTGGAAGTGGCGCAAGGAGGCCACCAAGTGAAG 1260
QY      1261 GACTGCAACGAGCGCGCAAGGCGCAACTTCTGGGCAAGATCTGGCCCAAGAGGCGCGC 1320
Db      1261 GACTGCAACGAGCGCGCAAGGCGCAACTTCTGGGCAAGATCTGGCCCAAGAGGCGCGC 1320
QY      1321 CCGGCAACTTCTGCAAGAACCGCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Db      1321 CCGGCAACTTCTGCAAGAACCGCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
QY      1381 CCGCGGAGAGCTTCGCTTGAAGAGACCAACCCCGCGCGCGCGCGCGCGCGCGCGCG 1440
Db      1381 CCGCGGAGAGCTTCGCTTGAAGAGACCAACCCCGCGCGCGCGCGCGCGCGCGCGCG 1440
QY      1441 CGGAGCCCTACCGGAGCGCGCTGACCGCGCTGCGAGCGCTGTCGAGCGCGCGCGCTG 1500
Db      1441 CGGAGCCCTACCGGAGCGCGCGCTGACCGCGCTGCGAGCGCGCGCGCGCGCGCTG 1500
```

```
Db      1441 CCGGAGCCCTACCGCGAGCGCCCTGACCGCGCTGCGAGCGCTGTTCGGCAGCGCGCGCTG 1500
QY      1501 AGCCAGTAA 1509
Db      1501 AGCCAGTAA 1509

RESULT 5
US-09-899-575-51
; Sequence 51, Application US/09899575
; Publication No. US20030223961A1
; GENERAL INFORMATION:
; APPLICANT: Zur Megede, Jan
; APPLICANT: Barnett, Susan W.
; APPLICANT: Egnelbrecht, Susan
; APPLICANT: van Rensburg, Estrelita Janse
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP01631.102
; CURRENT APPLICATION NUMBER: US/09/899,575
; CURRENT FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 09/475,704
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
; LENGTH: 1494
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HIV Type C
; US-09-899-575-51

Query Match      86.0%; Score 1297.6; DB 10; Length 1494;
Best Local Similarity 92.7%; Pred. No. 1.4e-271;
Matches 1402; Conservative 0; Mismatches 89; Indels 21; Gaps 3;

QY      1 ATGGGCGCGCGCGCCAGCATCTGCGCGCGGAGAAAGTGAAGTGGAGAAATCCGC 60
Db      1 ATGGGCGCGCGCGCCAGCATCTGAGCGGCGCAAGTGAAGTGGAGCGCATCCGC 60
QY      61 CTGCGCGCGCGCGCAAGAACCTACATGCTGAAGCACCTGTGTGGCGAGCGCGAG 120
Db      61 CTGCGCGCGCGCGCAAGAACCTACATGCTGAAGCACCTGTGTGGCGAGCGCGAG 120
QY      121 CTGAGGGCTTGCCTTGAACCCCGCGCTGCTGAGAGACCGCGAGGGCTGCAAGCATC 180
Db      121 CTGAGCGCTTGCCTTGAACCCCGCGCTGCTGAGAGACCGAGGGCTGCAAGCATC 180
QY      181 ATGAAGCAGCTGAGCGCGCGCTGCAAGACCGGCAACCGAGAGCTGCGAGCTGTACAAC 240
Db      181 ATCAAGCAGCTGAGCGCGCGCTGCAAGACCGGCAACCGAGAGCTGCGAGCTGTACAAC 240
QY      241 ACCGTGGCCACCTGTACTGTGCTGCAAGCGCGCATCGAGTCCGCGACCAAGAGGCC 300
Db      241 ACCGTGGCCACCTGTACTGTGCTGCAAGCGCGCATCGAGTCCGCGACCAAGAGGCC 300
QY      301 CTGCAACAAGATGAGGAGGAGCAACAAGTCCCAAGCAAGAACCCAGAGGCCCAAGAG 360
Db      301 CTGCAACAAGATGAGGAGGAGCAACAAGTCCCAAGCAAGAACCCAGAGGCCCAAGAG 360
QY      361 GCGGACGGCAAGGTGAGCCAGAACTACCCCATCTGTGCAAGACCTGCAAGGCGCGCATG 420
Db      361 GCGGACGGAAGGTGAGCCAGAACTACCCCATCTGTGCAAGACCTGCAAGGCGCGCATG 420
QY      421 CACCAAGGCATCAGCCCCCGCACCCCTGAAAGCTGGGTGAGGTGATCGAGAGAGGCC 480
Db      421 CACCAAGGCATCAGCCCCCGCACCCCTGAAAGCTGGGTGAGGTGATCGAGAGAGGCC 480
QY      481 TTCAGCCCGGAGGTGATCCCATGTTCAACCGCGCTGAGCGAGGGCGCCACCCCGCAGGAC 540
Db      481 TTCAGCCCGGAGGATCCCATGTTCAACCGCGCTGAGCGAGGGCGCCACCCCGCAGGAC 540
```

```
QY 541 CTGAACGATGTTGAACACCGTGGCGGCCACAGCGCCCATGAGATGCTGAAGAC 600
Db 541 CTGAACGATGTTGAACACCGTGGCGGCCACAGCGCCCATGAGATGCTGAAGAC 600
QY 601 ACCATCAACGAGAGCGCGCGAGTGGGACCGCTTGCAACCCCGTGACAGCGCGCGCGTG 660
Db 601 ACCATCAACGAGAGCGCGCGAGTGGGACCGCAACCCCGTGACAGCGCGCGCGTG 660
QY 661 GCGCGCGCGCGAGTGGCGGACCGCGCGCGAGCGACATCGCGCGCGCGCACAGCACCTG 720
Db 661 GCGCGCGCGCGAGTGGCGGACCGCGCGCGAGCGACATCGCGCGCGCGCACAGCACCTG 720
QY 721 CAGAGCAGATCGCTGGATGACCAAGCAACCCCGCTGCGCGCGCGACATCTAAG 780
Db 721 CAGAGCAGATCGCTGGATGACCAAGCAACCCCGCTGCGCGCGCGACATCTAAG 780
QY 781 CGGTGATCATCTGGGCTTGAACAAAGATCGTGGATGTACAGCCCGTGAGCATCTG 840
Db 781 CGGTGATCATCTGGGCTTGAACAAAGATCGTGGATGTACAGCCCGTGAGCATCTG 840
QY 841 GACATCCGCGAGGCGCGCGCGAGCGCTTCCGCGACTACGTGACCGCTTCTTAAGACC 900
Db 841 GACATCAAGCAGGCGCGCGCGAGCGCTTCCGCGACTACGTGACCGCTTCTTAAGACC 900
QY 901 CTGCGCGCGCGAGCAGCGCCACCCAGACGTGAAGAACTGGATGACCGAGACCTGTGGTG 960
Db 901 CTGCGCGCGCGAGCAGCGCCACCCAGACGTGAAGAACTGGATGACCGAGACCTGTGGTG 960
QY 961 CAGAACGCAACCCCGACTGCAAGACCATCTGCGCGCTCTGCGCGCGCGCGCACCTG 1020
Db 961 CAGAACGCAACCCCGACTGCAAGACCATCTGCGCGCTCTGCGCGCGCGCGCACCTG 1020
QY 1021 GAGAGATGATGACCGCTGCGAGGCGGTGGCGCGCGCGCGCACAGCGCGCGCTG 1080
Db 1021 GAGAGATGATGACCGCTGCGAGGCGGTGGCGCGCGCGCGCACAGCGCGCGCTG 1080
QY 1081 GCCGAGCGATGAGCGCCCAACAGCGTGAACATCATGATGAGAGAGCAACTTCAAG 1140
Db 1081 GCCGAGCGCATGAGCGCCCAACAGC--AAATCTGTGTGACGCGCAACTTCAAG 1137
QY 1141 GCGCGCGCGCAACGTCAAGTGTCAACTGCGCGCAAGAGGCGCAATGCGCAAGAAC 1200
Db 1138 GCGAGCAACCGCATCATCAAGTGTCAACTGCGCGCAAGTGGCGCAATGCGCGCAAC 1197
QY 1201 TGCGCGCGCGCGCGCAAGAGGCGTGGAGTGGCGCAAGAGGCGCACAGATGAAG 1260
Db 1198 TGCGCGCGCGCGCGCAAGAGGCGTGGAGTGGCGCAAGAGGCGCACAGATGAAG 1257
QY 1261 GACTGACCGAGCGCGCGCAACTTCTGGGCAAGATCTGCGCGCACAGAGGCGCG 1320
Db 1258 GACTGACCGAGCGCGCGCAACTTCTGGGCAAGATCTGCGCGCACAGAGGCGCG 1317
QY 1321 CCGCGCACTTCTGGAACCGCGAGCGAGC--CGCGCGCGCGCGCACCGTGCCACCGCC 1377
Db 1318 CCGCGCACTTCTGGAACCGCGCGAGCGCGCACCGCGCGCGCGCGCACCGCGCC 1377
QY 1378 CCGCGCGCGAGAGCTTCCGCTTGAAGAGACCAACCGCGCGCGCGCGCACAGAGCCCAAG 1437
Db 1378 CCGCGCGCGAGAGCTTCCGCTTGAAGAGACCAACCGCGCGCGCGCGCACAGAGGAG 1437
QY 1438 GACCGCGAGCGCTTACCGCGAGCGCGCTGACCGCGCGCAAGCTTTCGGCAGCGCGCC 1497
Db 1438 -----CGCGAGCGCGCTGACCAAGCTGAAGAGCTTTCGGCAGCGCGCC 1482
QY 1498 CTGAGCCAGTAA 1509
Db 1483 CTGAGCCAGTAA 1494
```

RESULT 6

US-09-899-575-99

; Sequence 99, Application US/09899575

```
; Publication No. US20030223961A1
; GENERAL INFORMATION:
; APPLICANT: Zur Megede, Jan
; APPLICANT: Barnett, Susan W.
; APPLICANT: Egnelbrecht, Susan
; APPLICANT: van Rensburg, Estrellita Janse
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP01631.102
; CURRENT APPLICATION NUMBER: US/09/899,575
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 09/475,704
; PRIOR FILING DATE: 1999-12-30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 99
; LENGTH: 1491
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Gag_TV2_C_ZAopt
US-09-899-575-99

Query Match      85.8%; Score 1294.8; DB 10; Length 1491;
Best Local Similarity 92.7%; Pred. No. 5.8e-271;
Matches 1401; Conservative 0; Mismatches 87; Indels 24; Gaps 3;

QY 1 ATGGGCGCGCGCGCGCGCGCATCTGCGCGCGCGAGAAAGTGGACAAGTGGAGAATCCGC 60
Db 1 ATGGGCGCGCGCGCGCGCGCATCTGCGCGCGCGAGCAAGTGGAGAATCCGC 60
QY 61 CTGCGCGCGCGCGCGCGCGCAAGAGCACTACATGCTGAAGCACTGTGTGGCGCGCGAG 120
Db 61 CTGCGCGCGCGCGCGCGCGCAAGAGCACTACATGCTGAAGCACTGTGTGGCGCGCGAG 120
QY 121 CTGAGGCGCTTCCCTGAAACCCCGCGCTGTGAGACCGCGCGAGGCGTGAAGCATC 180
Db 121 CTGAGCGCTTCCCTGAAACCCCGCGCTGTGAGACCGCGCGAGGCGTGAAGCATC 180
QY 181 ATGAAGCAGCTGAGCGCGCGCGCTGACAGCGCGCACCGAGAGCTGCGAGCTGTACAAC 240
Db 181 ATGAAGCAGCTGAGCGCGCGCGCGCTGACAGCGCGCACCGAGAGCTGCGAGCTGTACAAC 240
QY 241 ACCGTGGCCACCTGTACTGCGTGAACGCGCGCATCGAGTCCGCGACACCAAGAGGCC 300
Db 241 ACCGTGGCCACCTGTACTGCGTGAACGCGCGCATCGAGTCCGCGACACCAAGAGGCC 300
QY 301 CTGGAACAAGTGAAGAGAGAGCAACAAGTCCAGCAGAAAGCCAGAGCGCAAGAG 360
Db 301 CTGGAACAAGTGAAGAGAGAGCAACAAGTCCAGCAGAAAGCCAGAGCGCAAGAG 360
QY 361 GCGGACGCGCAAGTGAAGCAGAACTAACCCATCGTGCAGAACTGCAAGGCGCATGTG 420
Db 361 GCGGACGCGCAAGTGAAGCAGAACTAACCCATCGTGCAGAACTGCAAGGCGCATGTG 420
QY 421 CACCAAGGCGCATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
Db 421 CACCAAGGCGCATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
QY 481 TTCAAGCGCGCGAGTATCCCGCATGTTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
Db 481 TTCAAGCGCGCGAGTATCCCGCATGTTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
QY 541 CTGAACAGATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db 541 CTGAACAGATGTTGAACACCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 601 ACCATCAACGAGAGGCGCGCGAGTGGGACCGCGCTGCAACCCCGTGCAAGCGCGCGCGTG 660
Db 601 ACCATCAACGAGAGGCGCGCGAGTGGGACCGCGCTGCAACCCCGTGCAAGCGCGCGCGTG 660
QY 661 GCGCGCGCGCGAGTGGCGGACCGCGCGCGAGCGACATCGCGCGCGCGCACAGCACCTG 720
Db 661 GCGCGCGCGCGAGTGGCGGACCGCGCGCGAGCGACATCGCGCGCGCGCACAGCACCTG 720
```



```
Db 661 GCGGGGGCCAGATGCGCGAGGCCCCGCGGCGAGCATCGCCGGCAACCAAGACCCCTG 720
QY 721 CAGAGCAGATCGCTGGATGACCAAGACCCCGCTGCGCGGCGACATCTTCAAG 780
Db 721 CAGAGCAGATCGCTGGATGACCAAGACCCCGCTGCGCGGCGACATCTTCAAG 780
QY 781 CGGTGATCATCTGGGCTGAAACAAGATGTGCGGATGTACAGCCCGTGAGCATCTG 840
Db 781 CGGTGATCATCTGGGCTGAAACAAGATGTGCGGATGTACAGCCCGTGAGCATCTG 840
QY 841 GACATCCGCGAGGCGCCCAAGAGCCCTTCCGCGACTACGTGAGCCGCTTCTTCAAGAC 900
Db 841 GACATCAAGCAGGCGCCCAAGAGCCCTTCCGCGACTACGTGAGCCGCTTCTTCAAGAC 900
QY 901 CTGCGCGCGAGCAGAGCCCAAGAGCGTGAAGAACTGATGACCGAGACCTGTGCTG 960
Db 901 CTGCGCGCGAGCAGAGCCCAAGAGCGTGAAGAACTGATGACCGAGACCTGTGCTG 960
QY 961 CAGAACGCCAACCCCGCATGCAAGACCATCTGCGCGCTTCTGCGCCCGCGCACCTTG 1020
Db 961 CAGAACGCCAACCCCGCATGCAAGACCATCTGCGCGCTTCTGCGCCCGCGCACCTTG 1020
QY 1021 GAGAGATGATGACCGCTGCGAGGCGTGGCGCGCCCGGCAACAAGCCCGCGTGTG 1080
Db 1021 GAGAGATGATGACCGCTGCGAGGCGTGGCGCGCCCGGCAACAAGCCCGCGTGTG 1080
QY 1081 GCGAGGCGATGAGCGGCAAGCGCAACAGCGTGAACATGATGACAGAGAACTTCAAG 1140
Db 1081 GCGAGGCGATGAGCGGCAAGCGCAACAGCGTGAACATGATGACAGAGAACTTCAAG 1140
QY 1141 GCGCGCGCGCGCAAGCTCAAGTGTCTCAACTGCGGCAAGAGGCGCACATCGCAAGAAC 1200
Db 1141 GCGCGCGCGCGCGCGCGTGAAGTGTCTCAACTGCGGCGGCGAGGCGCACATCGCGCAAC 1200
QY 1201 TGCGCGCGCGCGCGCGCAAGAGGCTGTGGAAGTGGCGCAAGAGGCGCACCAATGAAG 1260
Db 1201 TGCGCGCGCGCGCGCGCAAGAGGCTGTGGAAGTGGCGCAAGAGGCGCACCAATGAAG 1260
QY 1261 GACTGCAACCGAGCGCGCAACTTCTTGGGCAAGATCTGCGCGCACCAAGGCGCGC 1320
Db 1261 GACTGCAACCGAGCGCGCAACTTCTTGGGCAAGATCTGCGCGCACCAAGGCGCGC 1320
QY 1321 CCGCGCAACTTCTTGCAGAACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
Db 1321 CCGCGCAACTTCTTGCAGAACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
QY 1378 CCGCGCGCGCGAGAGCTTCCGCTTGCAGAGAGAACCGCGCGCGCGCGCGCGCGCGCG 1437
Db 1381 CCGCGCGCGCGAGAGCTTCAAGTTCAAGGAGA-----CGCGCAAGAGAGCGCGCAAG 1431
QY 1438 GACCGCGAGCGCTTACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1497
Db 1432 G-----ACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1479
QY 1498 CTGAGCCAGTAA 1509
Db 1480 CTGAGCCAGTAA 1491

RESULT 7
US-09-967-464-63
; Sequence 63, Application US/09967464
; Publication No. US20030138453A1
; GENERAL INFORMATION:
; APPLICANT: O'Hagan, Derek
; APPLICANT: Otten, Gillis
; APPLICANT: Donnelly, John J.
; APPLICANT: Polo, John M.
; APPLICANT: Barnett, Susan
; APPLICANT: Singh, Mamohan
; APPLICANT: Ulmer, Jeffrey
; APPLICANT: Dubensky, Jr., Thomas W.
; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY OF HETEROLOGOUS NUCLEIC ACIDS
```

```
; FILE REFERENCE: PP16269.004
; CURRENT APPLICATION NUMBER: US/09/967,464
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/236,105
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/315,905
; PRIOR FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 63
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-09-967-464-63

Query Match      84.6%; Score 1276.8; DB 10; Length 1479;
Best Local Similarity 92.5%; Pred. No. 4.5e-267;
Matches 1399; Conservative 0; Mismatches 77; Indels 36; Gaps 4;

QY 1 ATGGGCGCGCGCGCGCGCATCTGCGCGCGGCAAGCTGGAAGTGAGAGATCCGC 60
Db 1 ATGGGCGCGCGCGCGCGCATCTGCGCGCGGCAAGCTGGAAGCTGGAAGCTGGAAGCTCCG 60
QY 61 CTGCGCGCGCGCGCGCAAGAGCACTACATGCTGAAGCAACCTGTGTGGGCGAGCGCGAG 120
Db 61 CTGCGCGCGCGCGCGCAAGAGTGTCTACATGATGAAGCAACCTGTGTGGGCGAGCGCGAG 120
QY 121 CTGAGGGCTTGCCTCTGAACCCCGCGCTGTGAGAACCGCGAGGCGTGCACAGATC 180
Db 121 CTGAGGAAGTTGCTCTGAACCCCGCGCTGTGAGAACCGAGAGGCGTGCACAGATC 180
QY 181 ATGAAGCAGCTGCAGCCCGCGCTGCAGAACCGGCAACGAGAGCTGCGAGCTGTACAAC 240
Db 181 ATCCGCGAGCTGCAGCCCGCGCTGCAGAACCGGCAACGAGAGCTGGAAGCTGTACAAC 240
QY 241 ACCGTGGCAACCTGTACTGCTGTCAGCGCGCGCATGAGGTCCGCAACCAAGAGGCC 300
Db 241 ACCGTGGCAACCTGTACTGCTGTCAGCGCGCGCATGAGGTCCGCAACCAAGAGGCC 300
QY 301 CTGGAACAAGATGAGAGAGAGCAACAAGTCCAGCAGAGAACCCAGAGGCCAAGAG 360
Db 301 CTGGAACAAGATGAGAGAGAGCAACAAGTCCAGCAGAGAACCCAGAGGCCAAGAG 360
QY 361 GCGGAC---GGCAAGGTGAGCGCAGAACTACCCCATCTGTGAGAACTGCAAGGCCAGATG 417
Db 361 GCGGACAAAGGCAAGGTGAGCGCAGAACTACCCCATCTGTGAGAACTGCAAGGCCAGATG 420
QY 418 GTGCACCAAGGCATCAGCCCGCGCACCTGAAAGCCTGTGGTGAAGGTATCGAGAGAAG 477
Db 421 GTGCACCAAGGCATCAGCCCGCGCACCTGAAAGCCTGTGGTGAAGGTATCGAGAGAAG 480
QY 478 GCCTTACGCGCGAGGTGATCCCATGTTCACCGCGCTGAGCGAGGCGCGCACCCCGCAG 537
Db 481 GCCTTACGCGCGAGGTGATCCCATGTTCACCGCGCTGAGCGAGGCGCGCACCCCGCAG 540
QY 538 GACCTGAACACGATGTTGAACAACCGTGGCGCGCACCAAGGCGCGCATGCAATGCTGAAG 597
Db 541 GACCTGAACACGATGTTGAACAACCGTGGCGCGCACCAAGGCGCGCATGCAATGCTGAAG 600
QY 598 GACACCATCAACGAGAGGCGCGCGAGTGGAACCGGCTGCAACCCCGTGCAAGCGCGCGCC 657
Db 601 GACACCATCAACGAGAGGCGCGCGAGTGGAACCGGCTGCAACCCCGTGCAAGCGCGCGCC 660
QY 658 GTGCGCGCGCGCGCGAGTGGCGCAACCCCGCGCGAGCGACATCGCGCGCGCACAGCAC 717
Db 661 ATGCGCGCGCGCGCGAGTGGCGAGCCCGCGCGAGCGACATCGCGCGCGCACAGCAC 720
QY 718 CTGCAGAGAGCAGATCGCTGATGACCAAGCAACCCCGCGCGCGTGGGCGACATCTAC 777
Db 721 CTGCAGAGAGCAGATCGCTGATGACCAAGCAACCCCGCGCGCGTGGGCGACATCTAC 780
QY 778 AAGCGGTGATCATCTGGGCGCTGAACAAGATCTGCGGATGTACAGCCCGTGAGCATC 837
```



```
Db      781 AAGCGGTGATCATCTGGGCTGAACAAGATCGTCCGATGTACAGCCCCGTGACATC 840
QY      838 CTGGACATCCGCAAGGCCCCCAAGAGCCCTTCCGCACTACGTGACCGCTTCTTCAAG 897
Db      841 CTGACATCAAGCAGGCCCCCAAGAGCCCTTCCGCACTACGTGACCGCTTCTTCAAG 900
QY      898 ACCCTGCGCGCCGAGAGGCCACCAAGACGTGAAGACTGATGACCGAGACCTGCTG 957
Db      901 ACCCTGCGCGCCGAGAGCACCAAGAGAGTGAAGACTGATGACCGAGACCTGCTG 960
QY      958 GTGAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCGCGCCACC 1017
Db      961 GTGAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCGCGCCAGC 1020
QY      1018 CTGAGAGAGATGATGACCGCTGCAAGGGCGGTGGCGGCCCGGCCACAAGCCCGCTG 1077
Db      1021 CTGAGAGAGATGATGACCGCTGCAAGGGCGGTGGCGGCCCGGCCACAAGCCCGCTG 1080
QY      1078 CTGCGCCGAGCGGATGAGCCAGGCCAACAGCGTGAACATCATGATGCAAGAGCACTTC 1137
Db      1081 CTGCGCCGAGCGGATGAGCCAGGCCAACACC--AGCGTATGATGCAAGAGCACTTC 1137
QY      1138 AAGGCCCCCGCGCAAGTCAAGTCTTCACTGCGGCAAGAGGGCCACATCGCCAAG 1197
Db      1138 AAGGCCCCCGCGCGCATCTGTCAAGTCTTCACTGCGGCAAGAGGGCCACATCGCCCCG 1197
QY      1198 AACTGCCCCGCCCCCGCAAGAGGGCTCTGGAAGTGGGCAAGAGGGCCACCATG 1257
Db      1198 AACTGCCCCGCCCCCGCAAGAGGGCTCTGGAAGTGGGCAAGAGGGCCACCATG 1257
QY      1258 AAGGACTGCAACCGAGCGCCCAAGGCTTCTGCGGCAAGATCTGCGCCCAAGAGGGC 1317
Db      1258 AAGGACTGCAACCGAGCGCCCAAGGCTTCTGCGGCAAGATCTGCGCCCAAGAGGGC 1317
QY      1318 CGCCCCGGCACTTCTGCAAGAACCGGAGCGGCCCGCCCAACCGTCCCAACCGCC 1377
Db      1318 CGCCCCGGCACTTCTGCAAGAGCCG-----CGCCGAGCCCAACCGCC 1359
QY      1378 CCCCCCGCGAGAGCTTCCGCTTCAAGGAGAGACCAACCCCGCCCAAGAGAGCCCAAG 1437
Db      1360 CCCCCCGCGAGAGCTTCCGCTTCAAGGAGAGACCAACCCCGCGCCAGAGAGCAAG 1419
QY      1438 GACCGCGAGAGCCCTTACCGCCCTGACCGCCCTGCGCAGACCTGTTCCGAGCGGCC 1497
Db      1420 G-----ACCGCGAGAGCCCTGACCAAGCTGAAGAGCCTGTTCCGCAACGACCCC 1467
QY      1498 CTGAGCCAGTAA 1509
Db      1468 CTGAGCCAGTAA 1479
```

RESULT 8

```
US-09-899-575-3
; Sequence 3, Application US/09899575
; Publication No. US20030223961A1
; GENERAL INFORMATION:
; APPLICANT: Zur Megede, Jan
; APPLICANT: Barnett, Susan W.
; APPLICANT: Egnelbrecht, Susan
; APPLICANT: van Rensburg, Estrelita Janse
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP01631.102
; CURRENT APPLICATION NUMBER: US/09/899,575
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 09/475,704
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Artificial Sequence
```

```
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic Gag
; OTHER INFORMATION: of HIV strain AF110965
US-09-899-575-3
```

```
Query Match      84.6%; Score 1276.8; DB 10; Length 1479;
Best Local Similarity 92.5%; Pred. No. 4.5e-267;
Matches 1399; Conservative 0; Mismatches 77; Indels 36; Gaps 4;
```

```
QY      1 ATGGGCGCCCGCGCCAGCATCTGCGCGCGGCAAGACTTGACAAGTGGAGAGATCCGC 60
Db      1 ATGGGCGCCCGCGCCAGCATCTGCGCGCGGCAAGCTGACCGCTGGAGCGCATCCGC 60
QY      61 CTGCGCCCCCGCGCAAGAGCACTACATGCTGAAGCACCTGTGTGGCCAGCCGCGAG 120
Db      61 CTGCGCCCCCGCGCAAGAGTGTACATGATGAAGCACCTGTGTGGCCAGCCGCGAG 120
QY      121 CTGAGGGCTTCCGCTTGAACCCCGGCTGTGAGACCGCGAGGGCTGCAAGCATC 180
Db      121 CTGAGGAAGTTCCGCTTGAACCCCGGCTGTGAGACCGAGAGGGCTGCAAGCATC 180
QY      181 ATGAAGAGCTGCAAGCCCGCTGCAAGCCGCAACGAGAGCTGCGAGCTGTACAAC 240
Db      181 ATCCGCCAGCTGCAACCCCGCTGCAAGCCGCAACGAGAGCTGAAGAGCTGTACAAC 240
QY      241 ACCGTGCCACCTGTACTGCTGTCAGCGCGGCATGAGGTCCGACACCAAGAGGCC 300
Db      241 ACCGTGCCACCTGTACTGCTGTCAGAGAGAGATGAGGTCCGACACCAAGAGGCC 300
QY      301 CTGGAACAAGATCGAGAGAGCAGAACAAGTCCAGCAGAGAACCAGAGGCCAAGAG 360
Db      301 CTGGAACAAGATCGAGAGAGCAGAACAAGTCCAGCAGAGATCCAGAGGCCAAGAG 360
QY      361 GCCGAC--GGCAAGGTGAGGCCAGAACTACCCATGCTGCAAGAACTGCAAGGCCAGATG 417
Db      361 GCCGACAAGGGCAAGGTGAGGCCAGAACTACCCATGCTGCAAGAACTGCAAGGCCAGATG 420
QY      418 GTGCAACGAGCCATCAGCCCCCGCAACCTGAAACGCTGGTGAAGGTGATCGAGAGAAG 477
Db      421 GTGCAACGAGCCATCAGCCCCCGCAACCTGAAACGCTGGTGAAGGTGATCGAGAGAAG 480
QY      478 GCCTTCAGCCCCGAGGTGATCCCATGTTTCAACCGCCCTGAGCGAGGGCGCAACCCCCAG 537
Db      481 GCCTTCAGCCCCGAGGTGATCCCATGTTTCAACCGCCCTGAGCGAGGGCGCAACCCCCAG 540
QY      538 GACCTGAACAGATGTTGAACAAGTGGGGGCCAACAGAGCGCCATGCAAGTGTGAAG 597
Db      541 GACCTGAACAGATGTTGAACAAGTGGGGGCCAACAGAGCGCCATGCAAGTGTGAAG 600
QY      598 GACACCATCAACGAGAGAGCGCGCGAGTGGGACCGGCTGCAACCCCGTGACGGCGGCC 657
Db      601 GACACCATCAACGAGAGAGCGCGCGAGTGGGACCGGCTGCAACCCCGTGACGGCGGCC 660
QY      658 GTGGCCCCCGCGCAGATGCGGCAACCCCGCGCGCAGGCAATCGCCGGCCACAGCACC 717
Db      661 ATCGCCCCCGCGCAGATGCGGCAAGCCCGCGCGCAGGCAATCGCCGGCCACAGCACC 720
QY      718 CTGCAAGAGCAGATCGCTGATGACCAAGAACCCCGCGCGCGCGCATCTAC 777
Db      721 CTGCAAGAGCAGATCGCTGATGACCAAGAACCCCGCGCGCGCATCTAC 780
QY      778 AAGCGGTGATCATCTGCGGCTGAACAAGATCGTGGGATGTACAGCCCCGTGAGCATC 837
Db      781 AAGCGGTGATCATCTGCGGCTGAACAAGATCGTGGGATGTACAGCCCCGTGAGCATC 840
QY      838 CTGGAACATCCGCGAAGGGCCCCCAAGAGCCCTTCCGCACTACAGTGAACGCTTCTTCAAG 897
Db      841 CTGGAACATCAAGCAGGCCCCCAAGAGCCCTTCCGCACTACAGTGAACGCTTCTTCAAG 900
QY      898 ACCCTGCGCGCGCGAGAGGCCACCAAGAGCTGAAGAACTGATGACCGAGACCTGCTG 957
Db      901 ACCCTGCGCGCGCGAGAGGCCACCAAGAGCTGAAGAACTGATGACCGAGACCTGCTG 960
```

QY	958	GTGCAGAACGGCAACCCCGACTGCAAGACCATCTGCGGCTCTGGCCCGGCGCCACC	1017
Db	961	GTGCAGAACGGCAACCCCGACTGCAAGACCATCTGCGGCTCTGGCCCGGCGCCAGC	1020
QY	1018	CTGAGAGAGATGATGACCGGCTGCGCAGGGCGTGGGCGGCCCGCCACAAAGCCGCGTG	1077
Db	1021	CTGAGAGAGATGATGACCGGCTGCGCAGGGCGTGGGCGGCCCGCCACAAAGCCGCGTG	1080
QY	1078	CTGGCCGAGCGCATGAGCCAGGCCAAACAGCGTGAAATCATCATGATGCAAGAGCAACTTC	1137
Db	1081	CTGGCCGAGCGCATGAGCCAGGCCAAACAC---AGCGTATGATGCAAGAGCAACTTC	1137
QY	1138	AAGGGCCCCCGCGCAACGTCAAGTGCTTCAACTGCGCAAGAGGGGCCACATCGCCAAG	1197
Db	1138	AAGGGCCCCCGCGCATGTCAAGTGCTTCAACTGCGCAAGAGGGGCCACATCGCCGC	1197
QY	1198	AACCTCGCGCCCCCGCAAGAGGGCTGTGGAAGTGCGGCAAGAGGGGCCACCATG	1257
Db	1198	AACCTCGCGCCCCCGCAAGAGGGCTGTGGAAGTGCGGCAAGAGGGGCCACCATG	1257
QY	1258	AAGCACTGCACCGAGCGCCAGGCCAACTTCTGTGGCAAGATCTGGCCCAAGGCGC	1317
Db	1258	AAGCACTGCACCGAGCGCCAGGCCAACTTCTGTGGCAAGATCTGGCCCAAGGCGC	1317
QY	1318	CGCCCCGGCACTTCTCTGCAGAACCGCAGCGAGCCCGGCCCAACCGTGCCCAACCGC	1377
Db	1318	CGCCCCGGCACTTCTCTGCAGAGCG-----CCCGAGCTCCACCGCC	1359
QY	1378	CCCCCCCCCGAGAGCTTCCGCTTTCAGAGAGACCACCCCCGCCCCAAGCAGAGGCCCAAG	1437
Db	1360	CCCCCCCCCGAGAGCTTCCGCTTTCAGAGAGACCACCCCCGGCCAAGAGCAGAGACCAAG	1419
QY	1438	GACCCGCGAGCCCTAACCGCGAGCCCTGACCGCCCTGCGCAGCTGTTTCGCGACGGCCCC	1497
Db	1420	G-----ACCGCGAGACCTGACCAAGCTGAAGAGCGCTGTTTCGCGCAAGACCCC	1467
QY	1498	CTGAGCCAGTAA	1509
Db	1468	CTGAGCCAGTAA	1479

```

RESULT 9
US-10-190-435-18
; Sequence 18, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 3162
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagRTmut_C
US-10-190-435-18

```

Query Match	84.5%	Score 1275.2;	DB 15;	Length 3162;
Best Local Similarity	92.5%;	Pred. No. 9.8e-267;		
Matches 1398; Conservative	0;	Mismatches 78;	Indels 36;	Gaps 4;
QY	1	ATGGGCGCCGCGCCAGCATCTGCGCGGAGAGACTGACAACTGGAGAGATCCGC	60	
DB	7	ATGGGCGCCGCGCCAGCATCTGCGCGGCGGAGAGACTGACCGCTGGAGCGCATCCGC	66	

QY	61	CTGCGCCCCGGGCGCAAGAGCACTACATGCTGAAGCACCTGTGTGGGCCAGCCGGCAG	120
Db	67	CTGCGCCCCGGGCGCAAGAGTGTACATGATGAAGCACCTGTGTGTGGGCCAGCCGGCAG	126
QY	121	CTGGAAGGGCTTTCGCCCTGTAAACCCCGGCTGTCTGGAAGACCGCCGAGGGCTGCACGATC	180
Db	127	CTGGAAGAGTTTCGCCCTGTAAACCCCGGCTGTCTGGAAGACCGCGAGGGCTGCACGATC	186
QY	181	ATGAAGCAGCTGCAACCCCGCCCTGCAGACCGGSCACCGAGAGCTGCGCAGCCTGTACAAC	240
Db	187	ATCCGCCAGCTGCACCCCGCCCTGCAGACCGGSCAGGAGAGCTGAAGAGCCTGTTCAAC	246
QY	241	ACCGTGGCCACCCCTGTACTGCGTGCACGCGCGCATCGAGGTTCGCGACCAAGAGGCC	300
Db	247	ACCGTGGCCACCCCTGTACTGCGTGCACGAGAAGATTCAGGTTCGCGACCAAGAGGCC	306
QY	301	CTGGAACAAGATCGAGGAGAGAGACAACAAGTCCAGCAGAGAACCAGCAGGCCAAGAG	360
Db	307	CTGGAACAAGATCGAGGAGAGAGACAACAAGTCCAGCAGAGAATCCAGCAGGCCAAGGCC	366
QY	361	GCCGAC--GGCAAGTGAAGCCAGAATAACCCCATCTGTGCAAGCCTGCAGAGCCAGATG	417
Db	367	GCCGACAAAGGCAAGTGAAGCCAGAATAACCCCATCTGTGCAAGCCTGCAGAGCCAGATG	426
QY	418	GTCACACCAAGGCATCAGCCCCCGCACCCCTGAACGCTTGGGTGAAGGTGATCGAGAGAAG	477
Db	427	GTCACACCAAGGCATCAGCCCCCGCACCCCTGAACGCTTGGGTGAAGGTGATCGAGAGAAG	486
QY	478	GCTTTCAGCCCCCGAAGTATCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCAG	537
Db	487	GCTTTCAGCCCCCGAAGTATCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCAG	546
QY	538	GACCTGAACAAGATGTTGAACAACCGTGGCGGSCACCAAGGCCCATGCAATGCTGAAG	597
Db	547	GACCTGAACAAGATGTTGAACAACCGTGGCGGSCACCAAGGCCCATGCAATGCTGAAG	606
QY	598	GACACCATCAACGAGGAGGCGCGCGAGTGGACCGCCTGCACCCCGTGCAGGCGGCCCC	657
Db	607	GACACCATCAACGAGGAGGCGCGCGAGTGGACCGCCTGCACCCCGTGCACGCGGCCCC	666
QY	658	GTCGCCCCCGGCGCAGATGCGCGACCCCGCGGCGACGCAATCGCCGCGGCCACAGCAC	717
Db	667	ATCGCCCCCGGCGCAGATGCGCGAGCCCGCGGCGACGCAATCGCCGCGGCCACAGCAC	726
QY	718	CTGCAGAGCAGATCGCTGATGACCAAGCAACCCCGCTGCGCCCGTGGCGCATCTAC	777
Db	727	CTGCAGAGCAGATCGCTGATGACCAAGCAACCCCGCATCCCGTGGCGCATCTAC	786
QY	778	AAGCGGTGATCATCTGGGCTGAACAAGATCGTGGATGTACAGCCCGTGAGCATC	837
Db	787	AAGCGGTGATCATCTGGGCTGAACAAGATCGTGGATGTACAGCCCGTGAGCATC	846
QY	838	CTGCACATCCGCGAGGCGCCCAAGAGCCCTTCGCGCATCACTAGTGAACCGCTTCAAG	897
Db	847	CTGCACATCAAGCAGGCGCCCAAGAGCCCTTCGCGCATCACTAGTGAACCGCTTCAAG	906
QY	898	ACCCTGCGCGCGAGCAGGCCAACCAAGACGTGAAGAACTGATGACCGAGACCTGCTG	957
Db	907	ACCCTGCGCGCGAGCAGGCCAACCAAGACGTGAAGAACTGATGACCGAGACCTGCTG	966
QY	958	GTCAGAAAGCCCAACCCCGCATGCAAGACCATCTGCGCGCTCTCGGCCCGGCCACAC	1017
Db	967	GTCAGAAAGCCCAACCCCGCATGCAAGACCATCTGCGCGCTCTCGGCCCGGCCAGC	1026
QY	1018	CTGAGAGAGATGACCGGCTGCCAGGGCGTGGGCGGCCCGGCCACAAGGCCGCGTG	1077
Db	1027	CTGAGAGAGATGACCGGCTGCCAGGGCGTGGGCGGCCCGGCCACAAGGCCGCGTG	1086
QY	1078	CTGGCCGAGGCGATGAGCCAGGCCCAACAGCGTGAACATCATGATGCAAGAGCAACTTC	1137
Db	1087	CTGGCCGAGGCGATGAGCCAGGCCCAACACCC--AGCGTGAATGATGCAAGAGCAACTTC	1143

QY	1078	1137
CTGCCCAGGCGATGAGCCCAAGCGTGAACATCATGTGACAGAAGCACTTC		
DB	1027	1086
CTGGAGGAGATGATGACCCGCTGCCAGGCGGTGGCGCGGCCCAAGCAAGGCCCGCGTG		
QY	1087	1143
CTGGCCGAGCGCATGAGCCAGGCCCAACACC---AGCGTGTGATGCAGAGCAAGCACTTC		

QY 1318 CGCCCCGCACTTCTTCAGAGAACCGCAGAGCCCGCCGCCCCCAACCGTGCCCAACCGCC 1377
|
Db 1324 CGCCCCGCACTTCTTCAGAGAGCGC-----CCCCGAGCCCAACCGCC 1365
|
QY 1378 CCCCCCGCAGAGCTTCGCTTCGAGAGAACCAACCCCGCCCCCAAGCAGAGCCCAAG 1437
|
Db 1366 CCCCCCGCAGAGCTTCGCTTCGAGAGAACCAACCCCGCCAGAGCAGAGAGCAAG 1425
|
QY 1438 GACCCGAGCCCTACCGCAGACCCCTGACCGCCCTGCGCAGCCTGTTGCGCAGCGGCC 1497
|
Db 1426 G-----ACCGGAGACCCCTGACAGCCTGAAGAGCCTGTTGCGCAACGACCCC 1473
|
QY 1498 CTGAGCCAGTAA 1509
|
Db 1474 CTGAGCCAGAAA 1485

RESULT 11

US-10-190-435-19
; Sequence 19, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Esterelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: P18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 4419
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagRTmutTatRevNef_C
US-10-190-435-19

Query Match 84.5%; Score 1275.2; DB 15; Length 4419;
Best Local Similarity 92.5%; Pred. No. 9.6e-267;
Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;

QY 1 ATGGGCGCCCGCGCCACATCTCGCGCGCGAGAGAGCTGAGCAAGTGGAGAGAGATCCGC 60
|
Db 7 ATGGGCGCCCGCGCCACATCTCGCGCGCGAGAGAGCTGAGCAAGTGGAGAGATCCGC 66
|
QY 61 CTGCGCGCGCGCGCGAGAGAGCACTACATGCTGAAGCACTGTGTGGGCCAGCCGCGAG 120
|
Db 67 CTGCGCGCGCGCGCGAGAGAGTGTACATGATGAAGCACTGTGTGTGGGCCAGCCGCGAG 126
|
QY 121 CTGAGAGGCTTCGCTCTGAACCCCGGCTGTGAGACCGCGAGGGCTGCAAGCAGATC 180
|
Db 127 CTGAGAGAGTTCGCTCTGAACCCCGGCTGTGAGACCGAGGGCTGCAAGCAGATC 186
|
QY 181 ATGAAGCAGCTGACGCGCGCGCTGAGACCGGACCGAGAGAGTGGCGAGCCTGTACAAC 240
|
Db 187 ATCCCGCAGCTGACCGCGCGCTGAGACCGGACCGAGAGAGTGAAGAGCCTGTACAAC 246
|
QY 241 ACCGTGGCCACCCCTGTACTGCGTGCAGCGCGGATGAGGTCCGCGACACCAAGAGAGCC 300
|
Db 247 ACCGTGGCCACCCCTGTACTGCGTGCAGAGAGATGAGGTCCGCGACACCAAGAGAGCC 306
|
QY 301 CTGACACAAGATCGAGAGAGAGCAAGACAAGTCCAGAGAAAGCCAGAGGCCCAAGAGAG 360
|
Db 307 CTGACACAAGATCGAGAGAGAGCAAGACAAGTCCAGAGAAAGTCCAGAGGCCCAAGAGCC 366
|
QY 361 GCGGAC---GGCAAGGTGAGCCAGAACTACCCCATCTGTGCAAGACCTGCAAGGCCAGATG 417
|
Db 367 GCGGACAAAGGCAAGGTGAGCCAGAACTACCCCATCTGTGCAAGACCTGCAAGGCCAGATG 426

QY 418 GTGACCAAGGCCATCAGCCCCCGCAACCTGAACGCTGGGTGAAGGTGATCGAGAGAAAG 477
|
Db 427 GTGACCAAGGCCATCAGCCCCCGCAACCTGAACGCTGGGTGAAGGTGATCGAGAGAAAG 486
|
QY 478 GCGTTACGCCCCGAGAGTGTATCCCAATGTTCAACCGCCCTGAGCGAGGGGCCCAACCCCAAG 537
|
Db 487 GCGTTACGCCCCGAGAGTGTATCCCAATGTTCAACCGCCCTGAGCGAGGGGCCCAACCCCAAG 546
|
QY 538 GACCTGAACACGATGTTGAACACCGTGGGGGCCCAACAGCCCGCTGACAGCGCGGCCCC 597
|
Db 547 GACCTGAACACGATGTTGAACACCGTGGGGGCCCAACAGCCCGCTGACAGCGCGGCCCC 606
|
QY 598 GACACCATCAACGAGAGAGCGCCCGGAGTGGGACCGCCCTGACAGCGCGGCCCC 657
|
Db 607 GACACCATCAACGAGAGAGCGCCCGGAGTGGGACCGCCCTGACAGCGCGGCCCC 666
|
QY 658 GTGGCCCCCGGCGAGATGCGCGACCCCGCGGAGCGACATCGCGCGGCCACAGCACC 717
|
Db 667 ATCGCCCCCGGCGAGATGCGCGAGCCCGCGGAGCGACATCGCGCGGCCACAGCACC 726
|
QY 718 CTGACAGAGCAGATCGCTGTGATGACCAACCCCGCGTGGCGGTGGCGCATCTAC 777
|
Db 727 CTGACAGAGCAGATCGCTGTGATGACCAACCCCGCATCCCGTGGCGCATCTAC 786
|
QY 778 AAGCGGTGATCATCTCGGCGCTGAACAAGATCGTGGATGACAGCCCGTGAGCATC 837
|
Db 787 AAGCGGTGATCATCTCGGCGCTGAACAAGATCGTGGATGACAGCCCGTGAGCATC 846
|
QY 838 CTGACATCCGCGAGAGCGCCCAAGAGACCTTCGCGACTACGTGAACCGCTTCTTCAAG 897
|
Db 847 CTGACATCAAGCAGAGGCGCCCAAGAGACCTTCGCGACTACGTGAACCGCTTCTTCAAG 906
|
QY 898 ACCCTGCGCGCGAGAGCGCACCGACGAGACGTGAAGACTGAACCGAGACCTGCTG 957
|
Db 907 ACCCTGCGCGCGAGAGCGCACCGACGAGAGCTGAAGACTGAACCGAGACCTGCTG 966
|
QY 958 GTGCAAGACGCCAACCCCGACTGCAAGACCATCTGCGGCTTCGCGCGCGGCCACCC 1017
|
Db 967 GTGCAAGACGCCAACCCCGACTGCAAGACCATCTGCGGCTTCGCGCGCGGCCACCC 1026
|
QY 1018 CTGAGAGAGATGACCGCTGCGAGGCGGTGGCGGCCCGCCACAAGGCCCGCGTG 1077
|
Db 1027 CTGAGAGAGATGACCGCTGCGAGGCGGTGGCGGCCCGCCACAAGGCCCGCGTG 1086
|
QY 1078 CTGCGCGAGGCGAGTGAAGCGAGCGCAACAGCGTGAACATCATGTGAGAGAGCACTTC 1137
|
Db 1087 CTGCGCGAGGCGAGTGAAGCGAGCGCAACAC---AGCGTATGATGAGAGAGCACTTC 1143
|
QY 1138 AAGGCGCGCGCGCAAGTCAAGTCTTCACTGCGGCAAGAGGGCCACATGCCCAAG 1197
|
Db 1144 AAGGCGCGCGCGCATGCTCAAGTCTTCACTGCGGCAAGAGGGCCACATGCCCGC 1203
|
QY 1198 AACTGCGCGCGCGCGCGCAAGAGGGCTGCTGAAGTGGGCAAGAGGGCCACAGATG 1257
|
Db 1204 AACTGCGCGCGCGCGCGCAAGAGGGCTGCTGAAGTGGGCAAGAGGGCCACAGATG 1263
|
QY 1258 AAGACTGCAACGAGCGCGCAACTTCTTGGGCAAGATCTGGCCAGCCACAAGGGC 1317
|
Db 1264 AAGACTGCAACGAGCGCGCAACTTCTTGGGCAAGATCTGGCCAGCCACAAGGGC 1323
|
QY 1318 CGCCCCGCACTTCTTCGAGAACCGCAGCGCGCGCCCGCCCAACCGTGCCCAACCGCC 1377
|
Db 1324 CGCCCCGCACTTCTTCGAGAGCGC-----CCCGAGCCCAACCGCC 1365
|
QY 1378 CCCCCCGCAGAGCTTCGCTTCGAGAGAACCAACCCCGCCCCCAAGCAGAGGCCCAAG 1437
|
Db 1366 CCCCCCGCAGAGCTTCGCTTCGAGAGAACCAACCCCGCCAGAGCAGAGAGCAAG 1425
|
QY 1438 GACCCGAGCCCTAACCGGAGCGCCCTGACCGCCCTGCGAGCCTGTCGAGCGCGGCC 1497
|
Db 1426 G-----ACCGGAGACCCCTGACAGCCTGAAGAGCCTGTTGCGCAACGACCCC 1473


```

; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: pp18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35
; LENGTH: 4483
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160mod.TV1.dv1v2-gagmod.BW65
US-10-190-435-35

```

```

Query Match      84.5%; Score 1275.2; DB 15; Length 4483;
Best Local Similarity 92.5%; Pred. No. 9.6e-267;
Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;

```

```

QY 1 ATGGGCGCGCGCGCAGCATCTGCGCGCGAGAGCTGAGCAAGTGAGAGAAAGATCCGC 60
Db 3005 ATGGGCGCGCGCGCAGCATCTGCGCGCGAGAGCTGAGCAAGTGAGAGAAAGATCCGC 3064

QY 61 CTGCGCGCGCGCGCGCAAGACATGCTGAGCACTGTGTGTGGCCAGCCGCGAG 120
Db 3065 CTGCGCGCGCGCGCGCAAGAGTGCTACATGATGAGCACTGTGTGTGGCCAGCCGCGAG 3124

QY 121 CTGAGGGCTTGTGCTGAAACCCGCGCTGTGAGACCGCGAGGCTGCAAGCAGATC 180
Db 3125 CTGAGAGTTGCTGCTGAAACCCGCGCTGTGAGACCGAGGCTGCAAGCAGATC 3184

QY 181 ATGAAGAGCTGACGCGCGCTGTGACGCGCACCGAGAGCTGCGAGCCTGTACAC 240
Db 3185 ATCGCGCAGCTGACCGCGCGCTGTGACGCGCACCGAGAGCTGAGAGCCTGTACAC 3244

QY 241 ACCGTGGCCACCCCTGTACTGCTGTCAGCGCGCATGAGTCCGCGACCAAGAGGCC 300
Db 3245 ACCGTGGCCACCCCTGTACTGCTGTCAGAGAGATCGAGTCCGCGACCAAGAGGCC 3304

QY 301 CTGACAAGATCGAGAGGAGCAGAACTCCAGCAGAAAGACCGCAGGCCAAGAG 360
Db 3305 CTGACAAGATCGAGAGGAGCAGAACTCCAGCAGAAAGATCCAGAGGCCAAGAGGCC 3364

QY 361 GCGGAC---GGCAAGGTGAGCCAGAACTACCCCATCTGTGAGAACTGTGAGGGCCAGATG 417
Db 3365 GCGGACAAAGGCAAGGTGAGCCAGAACTACCCCATCTGTGAGAACTGTGAGGGCCAGATG 3424

QY 418 GTGACACAGGCCATCAGCCCCCGCACCCCTGAACGCTGGGTGAAGGTGATCGAGAGAG 477
Db 3425 GTGACACAGGCCATCAGCCCCCGCACCCCTGAACGCTGGGTGAAGGTGATCGAGAGAG 3484

QY 478 GCCTTCAGCCCCGAGGTGATCCCATGTTCAACCGCCTGAGCGAGGGGCCACCCCCAG 537
Db 3485 GCCTTCAGCCCCGAGGTGATCCCATGTTCAACCGCCTGAGCGAGGGGCCACCCCCAG 3544

QY 538 GACCTGAACACGATGTTGAACACCGTGGGGGGCCACCAAGCCGCCCATGCAATGCTGAAG 597
Db 3545 GACCTGAACACGATGTTGAACACCGTGGGGGGCCACCAAGCCGCCCATGCAATGCTGAAG 3604

QY 598 GACACCATCAACGAGAGGCGCGCGAGTGGGACCGCTGACACCCCGTGCAGGCGCGCCCC 657
Db 3605 GACACCATCAACGAGAGGCGCGCGAGTGGGACCGCGTGCACCCCGTGCAGGCGCGCCCC 3664

QY 658 GTGCGCGCGCGCGCAGATGCGCGAGACCCCGCGCGCAGCCGACATGCGCGCGCCACCAAGACC 717
Db 3665 ATGCGCGCGCGCGCAGATGCGCGAGCGCGCGCGCAGCCGACATGCGCGCGCCACCAAGACC 3724

QY 718 CTGCAAGAGCAGATCGCTGTGATGACCAACAACCCCGCTGCGCGGTGGGGCAGATCTAC 777
Db 3725 CTGCAAGAGCAGATCGCTGTGATGACCAACAACCCCGCTGCGCGGTGGGGCAGATCTAC 3784

```

```

QY 778 AAGCGTGATCATCTGGGCTTGAAACAAGATCGTGGGATGTACAGCCCCGTGACATC 837
Db 3785 AAGCGTGATCATCTGGGCTTGAAACAAGATCGTGGGATGTACAGCCCCGTGACATC 3844

QY 838 CTGCATCCGCCAGGGGCCCAAGAGGCCCTTCCGCGACTACGTGAGCCGCTTCTTCAAG 897
Db 3845 CTGCATCAAGCAGGGGCCCAAGAGGCCCTTCCGCGACTACGTGAGCCGCTTCTTCAAG 3904

QY 898 ACCCTCGCGCGCGAGCAGGCCACCCAGACGTGAACAAGTGAATCCGAGACCCGCTCTG 957
Db 3905 ACCCTCGCGCGCGAGCAGAGCAAGCAACCCAGAGGTGAACAAGTGAATCCGAGACCCGCTCTG 3964

QY 958 GTGCAGAACGCCAACCCCGACTGTGAAGACCATCTGCGCGCTTGGCCCCCGGCCAAC 1017
Db 3965 GTGCAGAACGCCAACCCCGACTGTGAAGACCATCTGCGCGCTTGGCCCCCGGCCAAC 4024

QY 1018 CTGAGAGAGATGATGACCGCTGCGAGGGCGTGGGCGGGCCCCCGCCACAAGGCCCGCTG 1077
Db 4025 CTGAGAGAGATGATGACCGCTGCGAGGGCGTGGGCGGGCCCCCGCCACAAGGCCCGCTG 4084

QY 1078 CTGCGCGAGCGCATGAGCCAGGCCAACAGCGTGAACATCATGATGACAGAGCAACTTC 1137
Db 4085 CTGCGCGAGCGCATGAGCCAGGCCAACACC--AGCGTATGATGACAGAGCAACTTC 4141

QY 1138 AAGGCGCGCGCGCGCAAGCTCAAGTCTTCAACTGCGCGCAAGAGGGCCACATGCGCAAG 1197
Db 4142 AAGGCGCGCGCGCGCATGCTCAAGTCTTCAACTGCGCGCAAGAGGGCCACATGCGCGC 4201

QY 1198 AACTGCGCGCGCGCGCGCAAGAGGGCTGTGAAAGTGCGGCAAGAGGGCCACAGATG 1257
Db 4202 AACTGCGCGCGCGCGCGCAAGAGGGCTGTGAAAGTGCGGCAAGAGGGCCACAGATG 4261

QY 1258 AAGGACTGCAACGAGCGCCAGGCCAATTCTGTGGCAAGATCTGGCCAGCCACAAGGGC 1317
Db 4262 AAGGACTGCAACGAGCGCCAGGCCAATTCTGTGGCAAGATCTGGCCAGCCACAAGGGC 4321

QY 1318 CGCGCGCGCAACTTCTGTGAGAAACCGCAGCGAGCGCGCGCGCCACCGTGGCCACCGCC 1377
Db 4322 CGCGCGCGCAACTTCTGTGAGAGCG-----CGCGAGCCACCGCC 4363

QY 1378 CCGCGCGCGAGAGCTTCCGCTTGAAGAGACCAACCCCGCGCCCAAGCAGAGCCCAAG 1437
Db 4364 CCGCGCGCGAGAGCTTCCGCTTGAAGAGACCAACCCCGCGCCCAAGCAGAGCAAG 4423

QY 1438 GACCGCGAGCCCTACCGCAGCGCCCTGACCGCGCTGCGAGCCTGTTCGCAAGCGCCCC 1497
Db 4424 G-----ACCGGAGAGCCTGACCAAGCCTGAAGAGCCTGTTCGCAAGCAAGCCCC 4471

QY 1498 CTGAGCCAGTAA 1509
Db 4472 CTGAGCCAAATAA 4483

```

```

RESULT 14
US-10-190-435-34
; Sequence 34, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEDEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; FILE REFERENCE: pp18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 4606

```

```

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160mod.TV1.dv1-gagmod.BW965
US-10-190-435-34

```

```

Query Match      84.5%; Score 1275.2; DB 15; Length 4606;
Best Local Similarity 92.5%; Pred. No. 9.6e-267;
Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;

```

```

QY 1 ATGGGCGCCGCGCCAGCATCTGCGCGCGAGAAAGCTGACAAAGTGGAGAAAGATCCGC 60
Db 3128 ATGGGCGCCGCGCCAGCATCTGCGCGCGAGAAAGCTGACAAAGTGGAGAAAGATCCGC 3187
QY 61 CTGCGCCCGCGCGCGAGAAAGCACTAATGCTGAAGCACTGTGTGGAGCAGCCGCGAG 120
Db 3188 CTGCGCCCGCGCGCGAGAAAGTGTATGATGAAGCACTGTGTGGAGCAGCCGCGAG 3247
QY 121 CTGAGGAGCTTCCGCTGAACCCCGGCTGTGAGACCGCGAGGCTGCAAGCAGATC 180
Db 3248 CTGAGAAAGTTCCGCTGAACCCCGGCTGTGAGACCGAGGCTGCAAGCAGATC 3307
QY 181 ATGAAGCAGCTGACGCGCCCTGCAACCGGACCGAGAGCTGCGCAGCTGTACAAC 240
Db 3308 ATCCGCGCAGCTGACGCGCCCTGCAACCGGACCGAGAGCTGAAGAGCTGTACAAC 3367
QY 241 ACCGTGGCCACCTGTACTGCGTGACGCGCGCATCGAGTCCGCGACCAAGAGGCGC 300
Db 3368 ACCGTGGCCACCTGTACTGCGTGACGAGAAAGATCGAGTCCGCGACCAAGAGGCGC 3427
QY 301 CTGAGCAAGATCGAGAGAGAGCAAGCAAGTCCAGCAAGAACCGCAGGCGCAAGAG 360
Db 3428 CTGAGCAAGATCGAGAGAGAGCAAGCAAGTCCAGCAAGAACCGCAGGCGCAAGAG 3487
QY 361 GCCGAC--GGCAAGGTGAGCCAGAACTACCCCATCGTGAGAACTGAGGGCCAGATG 417
Db 3488 GCCGACAAAGGCAAGGTGAGCCAGAACTACCCCATCGTGAGAACTGAGGGCCAGATG 3547
QY 418 GTGACCAAGGCGCATCAGCCCCCGCACCTGAAAGCTGGTGAAGGTGATCGAGAGAG 477
Db 3548 GTGACCAAGGCGCATCAGCCCCCGCACCTGAAAGCTGGTGAAGGTGATCGAGAGAG 3607
QY 478 GCCTTCAGCCCCGAGGTGATCCCATGTTCAACGCGCTGAGCGAGGGGCCACCCGCCAG 537
Db 3608 GCCTTCAGCCCCGAGGTGATCCCATGTTCAACGCGCTGAGCGAGGGGCCACCCGCCAG 3667
QY 538 GACCTGAACACGATGTTGAACACCGTGGGGCGGCCACAGCGCGCCATGAGATGCTGAAG 597
Db 3668 GACCTGAACACGATGTTGAACACCGTGGGGCGGCCACAGCGCGCCATGAGATGCTGAAG 3727
QY 598 GACACCATCAACGAGAGGGCGCGCGAGTGGGAGCCGCTGACACCCCGTGACGGCGGCC 657
Db 3728 GACACCATCAACGAGAGGGCGCGCGAGTGGGAGCCGCGTGACACCCCGTGACGGCGGCC 3787
QY 658 GTGGCCCCCGCGCGAGATGCGCGAGCCCGCGCGAGCATCGCCGGCCACCAAGCACC 717
Db 3788 ATCGCCCCCGCGCGAGATGCGCGAGCCCGCGCGAGCATCGCCGGCCACCAAGCACC 3847
QY 718 CTGCAAGAGCAGATCGCTGTGATGACCAAGCAACCCCGCTGCCCGTGGGGGAGCATCTAC 777
Db 3848 CTGCAAGAGCAGATCGCTGTGATGACCAAGCAACCCCGCTGCCCGTGGGGGAGCATCTAC 3907
QY 778 AAGCGGTGATCATCTCTGGGCTGAACAAGATCGTGGGATGTACAGCCCGTGAGCATC 837
Db 3908 AAGCGGTGATCATCTCTGGGCTGAACAAGATCGTGGGATGTACAGCCCGTGAGCATC 3967
QY 838 CTGACATCCGCGAGGGCCCCAAGAGCCCTTCCGCGACTACGTGAGCCGCTTCTTCAAG 897
Db 3968 CTGACATCAAGCAGGGCCCCAAGAGCCCTTCCGCGACTACGTGAGCCGCTTCTTCAAG 4027
QY 898 ACCCTGCGCGCGAGAGGGCCACCGAGAGCTGAAGAACTGATGACCGAGACCCCTGCTG 957

```

```

Db 4028 ACCCTGCGCGCGAGAGCAGACCCCAAGAGGTGAAGAACTGATGACCGACACCTGCTG 4087
QY 958 GTGCAAGAACGCCAACCCCGACTGTCAAGAACCATCTTCCGCGCTCTCGGCCCGGCCACC 1017
Db 4088 GTGCAAGAACGCCAACCCCGACTGTCAAGAACCATCTTCCGCGCTCTCGGCCCGGCCACC 4147
QY 1018 CTGAGAGAGATGATGACCGCTGTCCAGAGGGCGGTGGGGCGGCCACCAAGGCCGCTG 1077
Db 4148 CTGAGAGAGATGATGACCGCTGTCCAGAGGGCGGTGGGGCGGCCACCAAGGCCGCTG 4207
QY 1078 CTGCGCGGAGCGGATGAGCCAGGCGCAACGCTGAACATCATGATGACAGAAAGCACTTC 1137
Db 4208 CTGCGCGGAGCGGATGAGCCAGGCGCAACGCTGAACATCATGATGACAGAAAGCACTTC 4264
QY 1138 AAGGAGCCCCCGCGCAACGTCAAGTCTTCAACTGCGGCAAGAGGGGCCACATCGCCAA 1197
Db 4265 AAGGAGCCCCCGCGCATCGTCAAGTCTTCAACTGCGGCAAGAGGGGCCACATCGCCCGC 4324
QY 1198 AACTGCGCGCGCCCCCGCAAGAGGGCTGTGAAAGTGCAGGCAAGAGGGGCCACAGATG 1257
Db 4325 AACTGCGCGCGCCCCCGCAAGAGGGCTGTGAAAGTGCAGGCAAGAGGGGCCACAGATG 4384
QY 1258 AAGGACTGACCGAGCGCGCAAGGCTTCTGTGGGCAAGATCTGGCCACGACCAAGGGC 1317
Db 4385 AAGGACTGACCGAGCGCGCAAGGCTTCTGTGGGCAAGATCTGGCCACGACCAAGGGC 4444
QY 1318 CGCCCCCGCAACTTCTCTGAGAAACCGCAGAGCGCGCGCCCCCAACCGTGGCCACCGCC 1377
Db 4445 CGCCCCCGCAACTTCTCTGAGAGCGG-----CGCCGAGCCACCGCC 4486
QY 1378 CCCCCCGCGAGAGCTTCCGCTTGAAGAGACCAACCCCGCCCCCAAGCAGAGGCCCAAG 1437
Db 4487 CCCCCCGCGAGAGCTTCCGCTTGAAGAGACCAACCCCGCCCCCAAGCAGAGGCCCAAG 4546
QY 1438 GACCGCGAGCCCTAACCGCGAGCGCTGACCGCGCTGCGCAGCTGTTGGGAGCGGCC 1497
Db 4547 G-----ACCGGAGACCTGACCGAGCTGAAAGAGCTGTTGGGAGCAGACCC 4594
QY 1498 CTGAGCCAGTAA 1509
Db 4595 CTGAGCCAGTAA 4606

```

```

RESULT 15
US-10-190-435-36
; Sequence 36, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; FILE REFERENCE: P18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36
; LENGTH: 4615
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: gp160mod.TV1.dv2-gagmod.BW965
US-10-190-435-36

```

```

Query Match      84.5%; Score 1275.2; DB 15; Length 4615;
Best Local Similarity 92.5%; Pred. No. 9.6e-267;
Matches 1398; Conservative 0; Mismatches 78; Indels 36; Gaps 4;

```

QY 1 ATGGCGCCCGCCGACGATCCTGCGCGCGAGAAAGCTGGACAAGTGGAGAGAAGATCCGC 60
 Db 3137 ATGGCGCCCGCCGCGCAGCATCTGCGCGCGCGCAAGCTGGACGCTGGAGCGCATCCGC 3196
 QY 61 CTGCGCCCGCGCGCGCAAGAGCACTACTGCTGAAGCACTGTGTGGGCCAGCCGCGAG 120
 Db 3197 CTGCGCCCGCGCGCGCAAGAGTGTCTACTGATGAAGCACTGTGTGGGCCAGCCGCGAG 3256
 QY 121 CTGAGAGGCTTGCCTCTGAACCCCGGCTGTGAGACCGCGAGGCTGCAAGCAGATC 180
 Db 3257 CTGAGAGAGTTTCCCTCTGAACCCCGGCTGTGAGACCAAGAGGCTGCAAGCAGATC 3316
 QY 181 ATGAAGCAGCTGCAGCGCGCGCTGCAAGCCGCAAGAGAGCTGCGCAGCTGTACAA 240
 Db 3317 ATCCCGCAGCTGCAGCGCGCGCTGCAAGCCGCAAGAGAGCTGAAGAGCTGTTCAA 3376
 QY 241 ACCGTGGCCACCTGTACTGCGTGACGCGCGCATCGAGGTCGCGCAACCAAGAGGCC 300
 Db 3377 ACCGTGGCCACCTGTACTGCGTGACGAGAGATCGAGGTCGCGCAACCAAGAGGCC 3436
 QY 301 CTGACAGAGATCGAGAGAGAGCAAGCAAGTCCAGAGAGAGAGCCAGAGGCCAAGAG 360
 Db 3437 CTGACAGAGATCGAGAGAGAGCAAGCAAGTCCAGAGAGAGATCCAGAGGCCAGGCC 3496
 QY 361 GCCGAC---GGCAAGGTAGCCAGAACTACCCCATCGTGAGAACTGCAAGGCCAGTG 417
 Db 3497 GCCGACAAAGGCGAAGGTAGAGCCAGAACTACCCCATCGTGAGAACTGCAAGGCCAGTG 3556
 QY 418 GTGCACAGGCGCATCAGCCCCCGCACCTGAACGCTGAGTGAAGGTGATCGAGAGAA 477
 Db 3557 GTGCACAGGCGCATCAGCCCCCGCACCTGAACGCTGAGTGAAGGTGATCGAGAGAA 3616
 QY 478 GCCTTCAGCCCCGAGGTGATCCCATGTTCAACCGCTGAGCGAGGCGCCACCCCGAG 537
 Db 3617 GCCTTCAGCCCCGAGGTGATCCCATGTTCAACCGCTGAGCGAGGCGCCACCCCGAG 3676
 QY 538 GACCTGAACACGATGTTGAACACCGTGGCGGCGCACAGGCCGCGCATGAGATGCTGA 597
 Db 3677 GACCTGAACACGATGTTGAACACCGTGGCGGCGCACAGGCCGCGCATGAGATGCTGA 3736
 QY 598 GACACCATCAACGAGAGAGCGCGCGAGTGGACCGCTGCACCCCGTGAGGCGCGGCC 657
 Db 3737 GACACCATCAACGAGAGAGCGCGCGAGTGGACCGCGTGACCCCGTGACGCGGCC 3796
 QY 658 GTGCGCCCGCGCGCAGATGCGGACCCCGCGGAGGACATCGCGCGCCACCAAGACC 717
 Db 3797 ATCGCCCCCGCGCAGATGCGGCGGAGCGCGGAGGACATCGCGCGCCACCAAGACC 3856
 QY 718 CTGACGAGCAGATCGCTGATGACGCAACCCCGCTGCGCGGCGGACATCTAC 777
 Db 3857 CTGACGAGCAGATCGCTGATGACGCAACCCCGCATCCCGTGGGCGACATCTAC 3916
 QY 778 AAGCGGTGATCATCTGGGCGCTGAACAAGATCGTGGGATGTACAGCCCGTGAGCATC 837
 Db 3917 AAGCGGTGATCATCTGGGCGCTGAACAAGATCGTGGGATGTACAGCCCGTGAGCATC 3976
 QY 838 CTGACATCCGCGAGGCGCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTCAAG 897
 Db 3977 CTGACATCAAGCAGGCGCCCAAGAGCCCTTCCGCGACTACGTGACCGCTTCTCAAG 4036
 QY 898 ACCCTGCGCGCGCGAGAGGCCCAACCAAGAGCTGAAGAACTGGATGACCGAGACCTGCTG 957
 Db 4037 ACCCTGCGCGCGCGAGAGGCCCAACCAAGAGCTGAAGAACTGGATGACCGAGACCTGCTG 4096
 QY 958 GTGAGAAAGCCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCGCCCGGCGCAAC 1017
 Db 4097 GTGAGAAAGCCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCGCCCGGCGCAAC 4156
 QY 1018 CTGAGAGAGATGATGACCGCTGCGAGGCGGTGGGCGGCCCGGCGCAAGGCCCGCGTG 1077
 Db 4157 CTGAGAGAGATGATGACCGCTGCGAGGCGGTGGGCGGCCCGGCGCAAGGCCCGCGTG 4216
 QY 1078 CTGCGCGAGCGATGAGCCAGGCCCAAGCGTGAACATCATGATGCAAGAGCAACTTC 1137

Db 4217 CTGGCCGAGGCGGATGAGCCAGGCCCAACCC---ACCGTATGATGACAGAGAGCACTTC 4273
 QY 1138 AAGGCGCCCGCGCGCAACGTCAGTGTCTTCAACTGCGGCAAGAGAGGCCACATCGCCAA 1197
 Db 4274 AAGGCGCCCGCGCGCATGTCAGTGTCTTCAACTGCGGCAAGAGAGGCCACATCGCCCG 4333
 QY 1198 AACTGCGCGCGCGCGCGCAAGAGGCTGTGAAGTGGCGCAAGAGGCGCCACAGATG 1257
 Db 4334 AACTGCGCGCGCGCGCGCAAGAGGCTGTGAAGTGGCGCAAGAGGCGCCACAGATG 4393
 QY 1258 AAGACTGCAACCGAGCGCCAGGCCAATTCTGTGGCAAGATCTGGCCCAAGAGGCC 1317
 Db 4394 AAGACTGCAACCGAGCGCCAGGCCAATTCTGTGGCAAGATCTGGCCCAAGAGGCC 4453
 QY 1318 CGCCCGCGCACTTCTTGCAAGACCGCAGCGAGCCCGCGCCCACTGTGCCCAACGCC 1377
 Db 4454 CGCCCGCGCACTTCTTGCAAGAGCG---CCCGAGGCCCAACGCC 4495
 QY 1378 CCCCCCGCGAGAGCTTCCGCTTGAAGAGACCAACCCCGCGCCCAAGAGAGGCCCAAG 1437
 Db 4496 CCCCCCGCGAGAGCTTCCGCTTGAAGAGACCAACCCCGCGCGCAAGAGAGAGCAAG 4555
 QY 1438 GACCGCGAGCCCTAACCGGAGCGCGCTGACCGCGCTGCGAGCGCTGTGGCAGCGGCC 1497
 Db 4556 G-----ACCGGAGACCTTGACAGCGCTGAAGAGCGCTGTGGCAAGCAAGCCCC 4603
 QY 1498 CTGAGCCCACTAA 1509
 Db 4604 CTGAGCCCACTAA 4615

Search completed: September 1, 2005, 23:17:25
 Job time : 943.802 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 1, 2005, 21:17.23 ; Search time 26.261 Seconds
(without alignments)
8578.900 Million cell updates/sec

Title: US-09-475-704A-4
Perfect score: 2840
Sequence: 1 atggcgccgcgcgcacgcat.....gcggccctgagccagtaa 1509

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_SPOOL_P/US09475704/runat_01092005_160751_19541/app_query.fasta_1.3342
-DB=Issued_Patents_AA -QFMT=faastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=Bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09475704_@CGN_1_1_53_@runat_01092005_160751_19541 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCtUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2457	86.5	492	4	US-09-991-258-5
2	2246	79.1	500	1	US-08-375-510-1
3	2246	79.1	500	2	US-08-487-657-1
4	2246	79.1	500	4	US-09-309-572-16
5	2246	79.1	500	4	US-09-718-096-16
6	2238	78.8	1350	4	US-09-952-060-35
7	2237	78.8	512	3	US-08-463-210-8
8	2237	78.8	512	3	US-09-124-900-2
9	2237	78.8	512	4	US-08-463-028-8
10	2223	78.3	500	2	US-08-816-155B-45
11	2223	78.3	500	3	US-09-079-587-45
12	2094	73.7	493	4	US-09-952-060-33

13	2018.5	71.1	478	1	US-08-127-499A-11	Sequence 11, Appl
14	2018.5	71.1	478	1	US-08-482-847-11	Sequence 11, Appl
15	2011	70.8	437	3	US-08-392-794A-2	Sequence 2, Appl1
16	1949	68.6	512	4	US-09-319-588C-4	Sequence 4, Appl1
17	1748.5	61.6	498	1	US-08-470-202-59	Sequence 59, Appl1
18	1748.5	61.6	498	1	US-08-471-770-59	Sequence 59, Appl1
19	1748.5	61.6	498	2	US-08-468-059-59	Sequence 59, Appl1
20	1748.5	61.6	498	3	US-09-109-916-59	Sequence 59, Appl1
21	1748.5	61.6	498	4	US-09-886-156-59	Sequence 59, Appl1
22	1748.5	61.6	498	4	US-09-886-149-59	Sequence 59, Appl1
23	1748.5	61.6	498	4	US-09-886-150-59	Sequence 59, Appl1
24	1748.5	61.6	498	4	US-09-886-159-59	Sequence 59, Appl1
25	1748.5	61.6	498	4	US-10-326-090-59	Sequence 59, Appl1
26	1726.5	60.8	498	1	US-08-470-202-60	Sequence 60, Appl1
27	1726.5	60.8	498	1	US-08-471-770-60	Sequence 60, Appl1
28	1726.5	60.8	498	2	US-08-468-059-60	Sequence 60, Appl1
29	1726.5	60.8	498	3	US-09-109-916-60	Sequence 60, Appl1
30	1726.5	60.8	498	4	US-09-886-156-60	Sequence 60, Appl1
31	1726.5	60.8	498	4	US-09-886-149-60	Sequence 60, Appl1
32	1726.5	60.8	498	4	US-09-886-150-60	Sequence 60, Appl1
33	1726.5	60.8	498	4	US-09-886-159-60	Sequence 60, Appl1
34	1726.5	60.8	498	4	US-10-326-090-60	Sequence 60, Appl1
35	1671.5	58.9	363	2	US-08-850-049-130	Sequence 130, App
36	1671.5	58.9	363	3	US-08-050-478-130	Sequence 130, App
37	1671.5	58.9	363	3	US-09-414-117-130	Sequence 130, App
38	1671.5	58.9	363	4	US-09-678-437-130	Sequence 130, App
39	1671.5	58.9	363	4	US-09-943-722-130	Sequence 130, App
40	1545.5	54.4	799	1	US-07-648-796A-7	Sequence 7, Appl1
41	1544	54.4	458	1	US-07-648-796A-5	Sequence 5, Appl1
42	1530.5	53.9	337	1	US-07-648-796A-1	Sequence 1, Appl1
43	1514.5	53.3	328	4	US-08-776-188C-77	Sequence 77, Appl1
44	1453.5	51.2	518	4	US-09-206-551-45	Sequence 45, Appl1
45	1410.5	49.7	294	3	US-09-370-368-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-09-991-258-5
; Sequence 5, Application US/09991258
; Patent No. 6783939
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Robert
; APPLICANT: Keith, Paula
; APPLICANT: Dryga, Sergey
; APPLICANT: Caley, Ian
; APPLICANT: Maughan, Maureen
; APPLICANT: Johnston, Robert
; APPLICANT: Davis, Nancy
; APPLICANT: Swanstrom, Ronald
; TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOMES WITH MODIFIED HIV GENES FOR USE
; TITLE OF INVENTION: VACCINES
; FILE REFERENCE: 0113.0001U3
; CURRENT APPLICATION NUMBER: US/09/991,258
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 09/902,537
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,995
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6783939e =
US-09-991-258-5
Alignment Scores:
Pred. No.: 2,32e-141 Length: 492
Score: 2457.00 Matches: 465

Percent Similarity: 95.22% Conservative: 13
 Best Local Similarity: 92.63% Mismatches: 14
 Query Match: 86.51% Indels: 10
 DB: 4 Gaps: 2
 US-09-475-704A-4 (1-1509) x US-09-991-258-5 (1-492)

QY 1 ATGGGGCCCGCGCCAGCATCTCGCGCGGAGAGCTGGAACAAGTGGAGAGATCCGC 60
 Db 1 MetAlaAlaArgAlaSerIleLeuArgGlyGluLysLeuAspLysTrpGluLysIleArg 20
 QY 61 CTGCGCCCGCGCGGCAAGAACACTATGCTGAAGCACTGTGTGGCCAGCCGCGAG 120
 Db 21 LeuArgProGlyGlyLysLysHisIleTyrMetLeuLysHisIleValTrpAlaSerArgGlu 40
 QY 121 CTGAGAGGCTTCGCGCTGAAACCCCGCGCTGTGAGACCGCGAGGCTGCAAGCATC 180
 Db 41 LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrSerGluGlyCysLysGlnIle 60
 QY 181 ATGAAGCAGCTGCAGCGCCCGCTGTGACAGCCGACCGAGAGCTGCGCAACCTGTACAC 240
 Db 61 MetLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluGluLeuLysSerLeuTyrAsn 80
 QY 241 ACCGTGCGCCACCTGTACTGCGTGACAGCCGCGCATCGAGTCCGCGCAACCAAGAGGCC 300
 Db 81 ThrValAlaThrLeuTyrCysValHisGluLysIleGluValArgAspThrLysGluAla 100
 QY 301 CTGACACAGATCGAGAGAGAGAGAACAGTCCACAGCAAGAACCCAGAGGCCAAGAG 360
 Db 101 LeuAspLysIleGluGluGlnGluAsnLysCysGlnGlnLysThrGlnGlnAlaLysAla 120
 QY 361 GCCGAGCGCAAGGTGAGCCAGAACTACCCCATCGTGAGAACCTGAGGCGCCAGATGGTG 420
 Db 121 AlaAspGlyLysValSerGlnAsnTyrProIleValGlnAsnLeuGlnGlyGlnMetVal 140
 QY 421 CACCAAGGCATCAGCCCGCGCACCCCTGAACGCGCTGGTGAAGGTGATCGAGAGAGGCC 480
 Db 141 HisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValIleGluGluLysAla 160
 QY 481 TTCAGCCCGGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGCGGCCACCCCGCAGAG 540
 Db 161 PheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGlnAsp 180
 QY 541 CTGAACACGATGTTGAACACCGTGGCGGCGCACCGCCGCGCATGAGATGCTGAAGAG 600
 Db 181 LeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLysAsp 200
 QY 601 ACCATCAACGAGAGGCGCGCGAGTGGGACCGCTGTGACCCCGTGAAGCGCGCGCCGTG 660
 Db 201 ThrIleAsnGluGluAlaAlaGluTrpAspArgLeuHisProValHisAlaGlyProIle 220
 QY 661 GCCCGCGCGCAGATGCGCGCACCCCGCGCGCAGCAGCATCGCGCGGCCACCGACCCCTG 720
 Db 221 AlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerThrLeu 240
 QY 721 CAGAGCAGATGCGCTGATGACCAACACCCCGCTGCGCCGCTGGGCGACATCTACAAG 780
 Db 241 GlnGluGlnIleAlaTrpMetThrSerAsnProProIleProValGlyAspIleTyrLys 260
 QY 781 CGGTGGATCATCTGGGCTGAACAAGATCGTGGGATGTACAGCCCGCTGAGCATCTCG 840
 Db 261 ArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProValSerIleLeu 280
 QY 841 GACATCCGCGCAGGCGCCCAAGAGAGCCCTTCGCGACTACGTGACCGCTTCTTCAAGACC 900
 Db 281 AspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhePheLysThr 300
 QY 901 CTGCGCGCGCGAGAGGCCACCGAGACGTGAAGACTGGATGACCGAGACCTGTGCTGTG 960
 Db 301 LeuArgAlaGluGlnAlaThrGlnGluValLysAsnTrpMetThrAspThrLeuLeuVal 320
 QY 961 CAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCGCGCGCGCCACCTCG 1020

Db 321 GlnAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeuGlyProGlyAlaThrLeu 340
 QY 1021 GAGAGATGATGACCGCTTCCCGAGGCGGTGGGCGGCGCCCGGCAAGGCCCGCTGTG 1080
 Db 341 GluGluMetMetThrAlaCysGlnGlnGlyValGlyGlyProGlyHisLysAlaArgValLeu 360
 QY 1081 GCCGAGCGGATGAGCCAGGCCACAGCGGTGAACATCATGATGCAAGAGGCCACATCGCCAAGAAC 1140
 Db 361 AlaGluAlaMetSerGlnThrAsnSerGlyAsnIleMetMetGlnArgSerAsnPheLys 380
 QY 1141 GCGCCCGCGCGCAACGTCAAGTCTTCAACTGCGGCAAGAGGCCACATCGCCAAGAAC 1200
 Db 381 GlyProArgArgIleValLysCysPheAsnCysGlyLysGluGlyHisIleAlaArgAsn 400
 QY 1201 TGCCGCGCGCGCGCAAGAGGCTGTGAACTGCGGCAAGAGGCCACCATGTAAG 1260
 Db 401 CysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGluGlyHisGlnMetLys 420
 QY 1261 GACTGCACCGAGGCGCCAGGCCACTTCTGGGCAAGATCTGGGCCAGCCACCAAGGCGCGC 1320
 Db 421 AspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSerHisLysGlyArg 440
 QY 1321 CCGGCGCACTTCTGTGCAAGACCGCAGCGAGCCCGCGCGCCACCGTGCACCGCGCCCC 1380
 Db 441 ProGlyAsnPheLeuGlnAsnArgProGlu-----ProThrAlaPro 454
 QY 1381 CCGGCGGAGAGCTTCCGCTTGGAGGAGACCAACCCCGCGCGCAAGCAGAGGCCCAAGGAC 1440
 Db 455 ProAlaGluSerPheArgPheGluGluThrThrProAlaProLysGlnGluProIleGlu 474
 QY 1441 CGCGAGCCCTACCGCGAGCGCCCTGACCGCGCTGTGGCGAGCCTGTTCGGCAGCGGCCCTG 1500
 Db 475 -----ArgGluProLeuThrSerLeuLysSerLeuPheGlySerAspProLeu 490
 QY 1501 AGCCAG 1506
 Db 491 SerGln 492

RESULT 2
 US-08-375-510-1
 ; Sequence 1, Application US/08375510
 ; Patent No. 5576421
 ; GENERAL INFORMATION:
 ; APPLICANT: Saito, Atsushi
 ; APPLICANT: Sinagawa, Hideo
 ; APPLICANT: Nakata, Atsuo
 ; TITLE OF INVENTION: HIV ANTIGEN
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Birch, Stewart, Kolasch and Birch
 ; STREET: P.O. Box 747
 ; CITY: Falls Church
 ; STATE: Virginia
 ; COUNTRY: U.S.A.
 ; ZIP: 22040-0747
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/375,510
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/985,949
 ; FILING DATE: 04-DEC-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Svensson, Leonard R.
 ; REGISTRATION NUMBER: 30,330
 ; REFERENCE/DOCKET NUMBER: 216-309P
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 241-1300

```

; TELEFAX: (703) 241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 500 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Human immunodeficiency virus type 1
US-08-375-510-1

```

Alignment Scores:	
Pred. No.:	1.42e-128
Score:	2246.00
Percent Similarity:	90.16%
Best Local Similarity:	83.66%
Query Match:	79.08%
DB:	1
	Length: 500
	Matches: 425
	Conservative: 33
	Mismatches: 36
	Indels: 14
	Gaps: 5

US-09-475-704A-4 (1-1509) X US-08-375-510-1 (1-500)

[illegible][illegible]

RESULT 3
 US-08-487-657-1
 ; Sequence 1, Application US/08487657
 ; Patent No. 5834267
 ;
 ; GENERAL INFORMATION:
 ;
 ; APPLICANT: Saito, Atsushi
 ; APPLICANT: Sinagawa, Hideo
 ; APPLICANT: Nakata, Atsuo
 ;
 ; TITLE OF INVENTION: HIV ANTIGEN
 ;
 ; NUMBER OF SEQUENCES: 4
 ;
 ; CORRESPONDENCE ADDRESS:
 ;
 ; ADDRESSEE: Birch, Stewart, Kolasch and Birch
 ;
 ; STREET: P.O. Box 747
 ;
 ; CITY: Falls Church
 ;
 ; STATE: Virginia
 ;
 ; COUNTRY: U.S.A.
 ;
 ; ZIP: 22040-0747
 ;

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,657
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/375,510
; FILING DATE: 18-JAN-1995
; APPLICATION NUMBER: US/07/985,949
; FILING DATE: 04-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Svensson, Leonard R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 216-309P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 241-1300
; TELEFAX: (703) 241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 500 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Human immunodeficiency virus type 1
US-08-487-657-1
```

Alignment Scores:

Pred. No.:	1.42e-128	Length:	500
Score:	2246.00	Matches:	425
Percent Similarity:	90.16%	Conservative:	33
Best Local Similarity:	83.66%	Mismatches:	36
Query Match:	79.08%	Indels:	14
DB:	2	Gaps:	5

US-09-475-704A-4 (1-1509) x US-08-487-657-1 (1-500)

```
QY 1 ATGGGCGCCCGCGCCAGCATCTGCGCGCGAGAGCTGGACAAGTGGAGAGAATCCGC 60
Db 1 MetGlyAlaArgAlaSerValLeuSerGlyGlyLeuAspLysTrpGluLysIleArg 20
QY 61 CTGGCCCCCGCGGCAAGACACTGATGCTGAAGCACTGTGTGGCCAGCCGCGAG 120
Db 21 LeuArgProGlyGlyLysLysGlnTyrLysLeuLysHisIleValTrpAlaSerArgGlu 40
QY 121 CTGGAGGGCTTGGCCCTGAACCCCGGCTGCTGAGACCGCGGCTGCAAGCATC 180
Db 41 LeuGluArgPheAlaValAsnProGlyLeuGluThrSerGluGlyCysArgGlnIle 60
QY 181 ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGACGAGAGCTGCGCAGCTGTACAAC 240
Db 61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGluGluLeuArgSerLeuTyrAsn 80
QY 241 ACCGTGGCCACCTGTACTGCGTGCAGCCCGGCATCGAGTCCGCGACACCAAGAGGCC 300
Db 81 ThrIleAlaValLeuTyrCysValHisGlnArgIleAspValLysAspThrLysGluAla 100
QY 301 CTGACAAGATCGAGAGGAGAGACAAGTCCCAAGACAAGACCAAGAGGCCAAGAG 360
Db 101 LeuAspLysIleGlyGluGlnAsnLysSerLysLysAlaGlnGlnAlaIleAla 120
QY 361 GCCGACGGC-----AAGGTGAGCCAGAATACTCCCATCTGTGCAAGAACTGCAAGGCC 411
Db 121 AspThrGlyAsnAsnSerGlnValSerGlnAsnTyrProIleValGlnAsnLeuGlnGly 140
```

```
QY 412 CAGATGTGCACAGAGCCCATCAGCCCCCGCACCTGAAAGCCTGGGTGAAGTGCAG 471
Db 141 GlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValGlu 160
QY 472 GAGAAGGCTTACGCCCGGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGCGCCACC 531
Db 161 GlyLysAlaPheSerProGluValIleProMetPheSerAlaLeuSerGluGlyAlaThr 180
QY 532 CCCCAAGACCTGAACACGATGTTGAACACCGTGGGCGGCGCACAGCGCCCATGCAGATG 591
Db 181 ProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaIleMetGlnMet 200
QY 592 CTGAAGACACCATCAACGAGAGGCGCGCCGAGCTGGAGCCGCTGCACCCCGTGCAGGCC 651
Db 201 LeuLysGluThrIleAsnGluGluAlaIleGluTrpAspArgLeuHisProValHisAla 220
QY 652 GGGCCCGTGGCCCCCGGCGCAGATGCGCGACCCCGCGGACGACATCGCCGCGCCACC 711
Db 221 GlyProIleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThr 240
QY 712 AGCACCTGCAGAGACAGATCGCCTGGATGACCAACACCCCGCTGCCGTGGCGGAC 771
Db 241 SerThrLeuGlnGlnIleGlyTyrMetThrHisAsnProIleProValGlyGlu 260
QY 772 ATCTACAAGCGGTGATCATCTGTGGCGCTGAACAAGATCGTGGCGATGTACAGCCCGTG 831
Db 261 IleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProThr 280
QY 832 AGCATCTGCAGATCCGCGGAGGCGCCCAAGAGACCCCTTCCGCGACTACGTGACCGCTTC 891
Db 281 SerIleLeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhe 300
QY 892 TTCAAGACCCCTGCGCGCGCGAGAGGCGCACCCAGACGTGAAGAACTGATGACGAGACC 951
Db 301 TyrLysThrLeuArgAlaGluGlnAlaSerGlnGluValLysAsnTrpMetThrGluThr 320
QY 952 CTGCTGTGTGCAGAACCCCAACCCCGACTGCAGAACCATCTGCGCGCTCTCGGCCCCGGC 1011
Db 321 LeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyProGly 340
QY 1012 GCCACCTTGAGAGATGATGACCGCTGCCAGGGCGGTGGGCGGCCCGGCGCAAGGCC 1071
Db 341 AlaThrLeuGluGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLysAla 360
QY 1072 CCGGTGCTGGCGGAGGGCGATGAGCCAG--GCCAACAGCGTGAACATCATGATGCAGAG 1128
Db 361 ArgValLeuAlaGluAlaMetSerGlnValThrAsnProAlaThrIleMetIleGlnLys 380
QY 1129 AGCAACTTCAAGGGCCCCCGGCGCAACGTCAAGTCTTCAATGCGGCAAGAGGGCCAC 1188
Db 381 GlyAsnPheArgAsnGlnArgLysThrValLysCysPheAsnCysGlyLysGluHis 400
QY 1189 ATGCCAAGAACTGCGCGCGCCCGCGCAAGAGGGCTGTGGAAGTGGCAAGAGAGGGC 1248
Db 401 IleAlaLysAsnCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGluGly 420
QY 1249 CACCAGATGAAGACTGCACCGAGCGCCAGGCCAACTTCTGGGCAAGATGCGCCAGC 1308
Db 421 HisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSer 440
QY 1309 CACAAGGGCGCGCGCGCAACTTCTGCAGAACCGCAGCGAGCCCGCGCCCAACCGTG 1368
Db 441 HisLysGlyArgProGlyAsnPheLeuGlnSerArgProGlu----- 454
QY 1369 CCCACGCGCGCGCGCGAGAGCTTCCGCTTC-----GAGAGACCAACCCCGCGCCCC 1422
Db 455 ProThrAlaProProGluGluSerPheArgPheGlyGluGluThrThrThrProSerGln 474
QY 1423 AAGCAGAGGCCCAAGAGCGGAGCCCTACCGCGAGGCCCTGACCGCCCTGCGCAGCCTG 1482
Db 475 LysGlnGluProIleAspLysGluLeuTyr-----ProLeuAlaSerLeuArgSerLeu 492
QY 1483 TTCGGCAGCGCGCCCTGAGCCAG 1506
```

Db 493 PheGlySerAspProSerSerGln 500

RESULT 4
US-09-309-572-16
; Sequence 16, Application US/09309572
; Patent No. 6440730
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral hybrid vectors pseudotyped with LCMV
; FILE REFERENCE: P50489
; CURRENT APPLICATION NUMBER: US/09/309,572
; CURRENT FILING DATE: 1999-05-11
; EARLIER APPLICATION NUMBER: DE 198 56 463
; EARLIER FILING DATE: 1998-11-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: gag polypeptide
US-09-309-572-16

Alignment Scores:
Pred. No.: 1.42e-128 Length: 500
Score: 2246.00 Matches: 425
Percent Similarity: 90.16% Conservative: 33
Best Local Similarity: 83.66% Mismatches: 36
Query Match: 79.08% Indels: 14
DB: 4 Gaps: 5

US-09-475-704A-4 (1-1509) x US-09-309-572-16 (1-500)

QY 1 ATGGGGCGCCGCGCCGACATCTCGCGGCGGAGAGCTGACAAAGTGGAGAGAGATCCGC 60
Db 1 MetGlyAlaArgAlaSerValLeuSerGlyGlyGluLeuAspLysTrpGluLysIleArg 20
QY 61 CTGCGCGCCGCGGCGGCAAGACACTGATGCTGAAGACCTGTGTGGCCAGCGCGGAG 120
Db 21 LeuArgProGlyGlyLysLysGlnTrpLysLeuLysHisIleValTrpAlaSerArgGlu 40
QY 121 CTGAGGGGCTTGGCCCTGAACCCCGGCTGTGAGAGCCGCGAGGCTGCAAGCAGATC 180
Db 41 LeuGluArgPheAlaValAsnProGlyLeuLeuGluThrSerGluGlyCysArgGlnIle 60
QY 181 ATGAAGCAGCTGACGCGCCGCTGACAGCCGCGACGAGAGAGCTGCGCAGCTGTACAAC 240
Db 61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGluGluLeuArgSerLeuTyrAsn 80
QY 241 ACCGTGGCCACCTGTACTGCGTGCAGCGCGGACGAGGTCCGCGACCAAGAGAGGCC 300
Db 81 ThrIleAlaValLeuTyrCysValHisGlnArgIleAspValLysAspThrLysGluAla 100
QY 301 CTGACAAGATCGAGAGAGAGAGAGAAAGTCCCAAGAGAGAGAGAGAGAGAGAGAG 360
Db 101 LeuAspLysIleGluGluGlnGlnAsnLysSerLysLysAlaGlnGlnAlaAla 120
QY 361 GCCGACGCGC-----AAGGTGAGCGCAAGACTACCCCATCGTGCAGAACTGACGGGC 411
Db 121 AspThrGlyAsnAsnSerGlnValSerGlnAsnTyrProIleValGlnAsnLeuGlnGly 140
QY 412 CAGATGGTGCACACGAGGACATCAGCCCGCGACCTGAAGCGCTGGTGAAGGTGATCGAG 471
Db 141 GlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValValGlu 160
QY 472 GAGAAAGCCTTCAAGCCCGAGAGGTGATCCCATGTTCAACCGCCCTGAGAGAGGCGCCAC 531
Db 161 GluLysAlaPheSerProGluValIleProMetPheSerAlaLeuSerGluGlyAlaThr 180
QY 532 CCCAGGACCTGAACACGATGTTGAACACCGTGGCGCGCCACGAGCGCCCATGACGATG 591
|||||

Db 181 ProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMet 200
QY 592 CTGAAGGACACCATCAACGAGAGAGGCGCGGAGTGGAGCCGCTGCACCCCGTGCAGGCC 651
Db 201 LeuLysGluThrIleAsnGluGluAlaAlaGluTrpAspArgLeuHisProValHisAla 220
QY 652 GGGCCCGTGGCGCGCGCGAGATGCGCGAGCCCGCGGAGCGGACATCGCGCGGCCACC 711
Db 221 GlyProIleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThr 240
QY 712 AGCACCTGACAGACAGATCGCTGATGACCAACACCCCGCTGCGCGCGCGGAC 771
Db 241 SerThrLeuGlnGluGlnIleGlyTrpMetThrHisAsnProIleProValGlyGlu 260
QY 772 ATCTACAAGCGGTGATCATCTGCGGCTGAAACAAGATCGTGGAGTGTACAGCGCGTG 831
Db 261 IleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProThr 280
QY 832 AGCATCTGACATCCGCGAGGCGCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTC 891
Db 281 SerIleLeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhe 300
QY 892 TTCAGACCCCTGCGCGCGCGAGAGGCGCACCGACGAGACGTAAGAAGTGTGACCGAGACC 951
Db 301 TyrLysThrLeuArgAlaGluGlnAlaSerGlnGluValLysAsnTrpMetThrGluThr 320
QY 952 CTGCTGTGTGACAAAGCCCAACCCCGACTGCAAGACCATCTCGCGCTCTCGGCGCGGC 1011
Db 321 LeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyProGly 340
QY 1012 GCCACCTGAGAGAGATGATGACCGCTGCCAGGCGGTGGCGCGCGCGCGCAAGGCC 1071
Db 341 AlaThrLeuGluGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLysAla 360
QY 1072 CGCGTGTGCGCGAGCGGATGAGCCAG---GCCAAGCGGTGAACATCATGATGACAGAG 1128
Db 361 ArgValLeuAlaGluAlaMetSerGlnValThrAsnProAlaThrIleMetIleGlnLys 380
QY 1129 AGCAACTTCAAGGCGCGCGCGCAACGTCAAGTCTTCAACTCGCGCAAGAGGCGCAC 1188
Db 381 GlyAsnPheArgAsnGlnArgLysThrValLysCysPheAsnLysGlyLysGluGlnHis 400
QY 1189 ATGCCCAAGAACTGCCGCGCGCGCGCAAGAGAGGCTGTGAAGTGGCGCAAGAGGCGC 1248
Db 401 IleAlaLysAsnLysAlaArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGluGly 420
QY 1249 CACCATGTAAGGACTGCACCGCGCGCGCAACTTCTGCGGCAAGATCTGCGCCAGC 1308
Db 421 HisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSer 440
QY 1309 CACAAGGCGCGCGCGCGCAACTTCTGCAAGACCGCGAGCGCGCGCGCGCGCGCGT 1368
Db 441 HisLysGlyArgProGlyAsnPheLeuGlnSerArgProGlu----- 454
QY 1369 CCCACCGCGCGCGCGCGCGAGAGCTTCCGCTTC-----GAGAGACCAACCCCGCGCCC 1422
Db 455 ProThrAlaProProGluGluSerPheArgPheGlyGluGluThrThrThrProSerGln 474
QY 1423 AAGCAGAGAGCCCAAGAGACCGCGAGCCCTACCGCGAGCGCGCGCGCGCGCGCTG 1482
Db 475 LysGlnGluProIleAspLysGluLeuTyr-----ProLeuAlaSerLeuArgSerLeu 492
QY 1483 TTCGCGACGCGCGCGCGCTGAGCCAG 1506
Db 493 PheGlySerAspProSerSerGln 500

RESULT 5
US-09-718-096-16
; Sequence 16, Application US/09718096
; Patent No. 6589763
; GENERAL INFORMATION:
; APPLICANT: Von Jaer, Meike-Dorothee
; TITLE OF INVENTION: RETROVIRAL HYBRID VECTORS PSEUDOTYPED WITH LCMV

```
; FILE REFERENCE: 35-195
; CURRENT APPLICATION NUMBER: US/09/718, 096
; CURRENT FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: DE 19856463.5
; PRIOR FILING DATE: 1998-11-26
; PRIOR APPLICATION NUMBER: EP 99250415.9
; PRIOR FILING DATE: 1999-11-25
; PRIOR APPLICATION NUMBER: US 09/309,572
; PRIOR FILING DATE: 1999-05-11
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: gag polyprotein
US-09-718-096-16
```

```
Alignment Scores:
Pred. No.: 1.42e-128 Length: 500
Score: 2246.00 Matches: 425
Percent Similarity: 90.16% Conservative: 33
Best Local Similarity: 83.66% Mismatches: 36
Query Match: 79.08% Indels: 14
DB: 4 Gaps: 5
```

US-09-475-704A-4 (1-1509) x US-09-718-096-16 (1-500)

```
QY 1 ATGGGCGCCGCGCGCAGCATCTGCGCGCGAGAAAGCTGACAAGTGGAGAGATCCGC 60
Db 1 MetGlyAlaArgAlaSerValLeuSerGlyGlyLeuAspLysTrpGlyLysIleArg 20
QY 61 CTGCGCGCCGCGCGCAGAAAGCACTACATGCTGAAGCACTGTGGCCAGCCGCGAG 120
Db 21 LeuArgProGlyGlyLysLysGlnTyrLysLeuLysHisIleValTrpAlaSerArgGlu 40
QY 121 CTGAGGGCTTTCGCGCTGAACCCCGCGCTGCTGAGACCGCGCGAGGCTGCAAGCATC 180
Db 41 LeuGluArgPheAlaValAsnProGlyLeuLeuGluThrSerGluLysCysArgGlnIle 60
QY 181 ATGAAGCAGCTGCAGCCCGCGCTGCAGACCGGCACCGAGAGCTGCAGCCTGTACAAC 240
Db 61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGluLysLeuArgSerLeuTyrAsn 80
QY 241 ACCGTGGCACCCCTGTAAGTGCAGCCCGCATGCAGAGTCCGCGACCAAGAGGCC 300
Db 81 ThrIleAlaValLeuTyrCysValHisGlnArgIleAspValLysAspThrLysGluAla 100
QY 301 CTGACAAAGATCGAGAGAGAGAGAGAAAGTCCCGACAGAGAGAGAGAGAGAGAG 360
Db 101 LeuAspLysIleGluGluGluGlnAsnLysSerLysLysAlaGlnGlnAlaAla 120
QY 361 GCCGACGGC-----AAGGTGAGCGCAACTACCCCATCTGTGCAGAACTGCAGGCC 411
Db 121 AspThrGlyAsnAsnSerGlnValSerGlnAsnTyrProIleValGlnAsnLeuGlnGly 140
QY 412 CAGATGTCACACGAGCCATCAGCCCGCGACACCTGAACGCTGGGTGAAGTGATCAG 471
Db 141 GlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaIleValLysValGlu 160
QY 472 GAGAAGCCTTCAGCCCGCAGGTGATCCCATATGTTCAACCGCCCTGAGCGAGGCCACC 531
Db 161 GluLysAlaPheSerProGluValIleProMetPheSerAlaLeuSerGluGlyAlaThr 180
QY 532 CCCAAGGACCTGAACACAGATGTGAACACCGTGGCGCGCCACAGGCCCATGCAGATG 591
Db 181 ProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMet 200
QY 592 CTGAAGGACACCATCAACGAGAGAGCGCCGAGTGGAGCGCCTGCACCCCGTGCAGGCC 651
Db 201 LeuLysGluThrIleAsnGluGluAlaAlaGluTrpAspArgLeuHisProValHisAla 220
```

```
QY 652 GGGCCCGTGCGCCCGCGCCAGATGCGCGACCCCGCGGAGAGCATCGCCGCGCCACC 711
Db 221 GlyProIleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThr 240
QY 712 AGCACCCCTGCAGAGACAGATCGCGCTGGATGACCAAGCAACCCCGCTGCCGCGCGAC 771
Db 241 SerThrLeuGlnGlnGlnIleGlyTyrMetThrHisAsnProProIleProValGlyGlu 260
QY 772 ATCTACAAGCGGTGATCATCTCTGGGCGCTGAACAAGATCGTGGCGATGTACAGCCCGTG 831
Db 261 IleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProThr 280
QY 832 AGCATCTGCACATCCGCCAGAGCGCCCAAGAGCCCTTCGCGACTACGTGACCGCTTC 891
Db 281 SerIleLeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhe 300
QY 892 TTCAAGACCCCTGCGCGCGCGAGAGAGCCACCCAGACGCTGAAGAACTGATGACCGAGACC 951
Db 301 TyrLysThrLeuArgAlaGluGlnAlaSerGlnGluValLysAsnTrpMetThrGluThr 320
QY 952 CTGCTGTGTGCAGAAAGCCCAACCCCGACTGCAGAACCATCTCTGCGCGCTCTCGCGCCGCG 1011
Db 321 LeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyProGly 340
QY 1012 GCCACCCCTGAGAGATGATGATGACCGCTGCAGAGGCGCTGGCGGCGCCCGCGCAAGGCC 1071
Db 341 AlaThrLeuGluGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLysAla 360
QY 1072 CGCGTCTGCGCGAGGCGAGTGAAGCCAG--GCCAAGACGCTGAACATCATGATGCAGAAAG 1128
Db 361 ArgValIleuAlaGluAlaMetSerGlnValThrAsnProAlaThrIleMetIleGlnLys 380
QY 1129 AGCAACTTCAAGGGCCCCCGCGCGCAACGTCAAGTCTTCAACTGCGGCAAGAGGGCCAC 1188
Db 381 GlyAsnPheArgAsnGlnArgLysThrValLysCysPheAsnLysGlyLysGluGlyHis 400
QY 1189 ATGCGCAAGAACTGCGCGCGCGCGCGCGCAAGAGGGCTGCTGAAGTGGCGCAAGAGGGGC 1248
Db 401 IleAlaLysAsnLysCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGluGly 420
QY 1249 CACCAGATGAAGGACTGCACCGAGCGCGCAAGCTTCTGCGGCAAGATCTGCGCCAGC 1308
Db 421 HisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSer 440
QY 1309 CACAAGGGCGCGCGCGCAACTTCTGCAAGAACCGCAGCGAGCGCGCGCGCCACCGTG 1368
Db 441 HisLysGlyArgProGlyAsnPheLeuGlnSerArgProGlu----- 454
QY 1369 CCCACGCGCGCGCGCGCGAGAGCTTCGCGCTTC-----GAGAGACCAACCGCGCGCCC 1422
Db 455 ProThrAlaProProGluGluSerPheArgPheGlyGluGluThrThrThrProSerGln 474
QY 1423 AAGCAGAGGCCCAAGAGCGCGAGCCCTACCGCGAGCGCCCTGACCGCCCTGCGCAGCCTG 1482
Db 475 LysGlnGluProIleAspLysGluLeuTyr-----ProLeuAlaSerLeuArgSerLeu 492
QY 1483 TTCGCGAGCGCGCCCTGAGCCAG 1506
Db 493 PheGlySerAspProSerSerGln 500
```

RESULT 6
US-09-952-060-35
; Sequence 35, Application US/09952060
; Patent No. 6733993
; GENERAL INFORMATION:
; APPLICANT: Emimi, Emilio A.
; APPLICANT: Youil, Rima
; APPLICANT: Bett, Andrew J.
; APPLICANT: Chen, Ling
; APPLICANT: Kaslow, David C.
; APPLICANT: Shiver, John W.
; APPLICANT: Toner, Timothy J.
; APPLICANT: Casimiro, Danilo R.


```
/ TITLE OF INVENTION: ENHANCED FIRST GENERATION ADENOVIRUS
/ TITLE OF INVENTION: VACCINES EXPRESSING CODON OPTIMIZED HIV1-GAG, POL, NEF AND
/ TITLE OF INVENTION: MODIFICATIONS
/ FILE REFERENCE: 20747Y
/ CURRENT APPLICATION NUMBER: US/09/952, 060
/ CURRENT FILING DATE: 2001-09-14
/ PRIOR APPLICATION NUMBER: PCT/US01/28861
/ PRIOR FILING DATE: 2001-09-14
/ PRIOR APPLICATION NUMBER: 60/317,814
/ PRIOR FILING DATE: 2001-09-07
/ PRIOR APPLICATION NUMBER: 60/279,056
/ PRIOR FILING DATE: 2001-03-27
/ PRIOR APPLICATION NUMBER: 60/233,180
/ PRIOR FILING DATE: 2000-09-15
/ NUMBER OF SEQ ID NOS: 38
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 35
/ LENGTH: 1350
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Codon optimized gag-1A pol fusion
US-09-952-060-35

Alignment Scores:
Pred. No.: 5,01e-128 Length: 1350
Score: 2238.00 Matches: 423
Percent Similarity: 90.35% Conservative: 36
Best Local Similarity: 83.27% Mismatches: 35
Query Match: 78.80% Indels: 14
DB: 4 Gaps: 5

US-09-475-704A-4 (1-1509) x US-09-952-060-35 (1-1350)
QY 1 ATGGCGCGCCGCGCCACATCTCGCGCGCGAGAGCTGGACAAGTGGAGAGAAGATCCGC 60
Db 1 MetGlyAlaArgAlaSerValLeuSerGlyGlyGluLeuAspLysTrpGluLysIleArg 20
QY 61 CTGCGCCCCGCGCGGAGAGACACTACATGCTGAAGCACTGTGTGGCCAGCCGCGAG 120
Db 21 LeuArgProGlyGlyLysLysLysTyrlsLysLeuYsHisIleValITrpAlaSerArgGlu 40
QY 121 CTGAGGGGCTTCGCCCTGAACCCCGGCTGTGAGACCCGCGAGGGCTGCAAGCAGATC 180
Db 41 LeuGluArgPheAlaValAsnProGlyLeuLeuGluThrSerGluGlyCysArgGlnIle 60
QY 181 ATGAAGCAGCTGCAGCGCCCGCTGCAGACCGGACCGAGAGAGCTGCCAGCCTGTACAC 240
Db 61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGluGluLeuArgSerLeuTyraSn 80
QY 241 ACCGTGGCCACCTGTACTGCTGTCAGCGCGGATCGAGGTCCGCGACACCAAGAGGCC 300
Db 81 ThrValAlaThrLeuTyrcysValHisGlnLysIleAspValLysAspThrLysGluAla 100
QY 301 CTGACACAGATCGAGAGGAGCAGAAAGTCCACAGCAGAAAGCCAGCAGGCCAAGGAG 360
Db 101 LeuGluLysIleGluGluGluGlnAsnLysSerLysLysAlaGlnGlnAlaAla 120
QY 361 GCCGACGGC-----AAGGTAGCCAGAACTACCCCATGCTGTCAGAACTGTCAGGGC 411
Db 121 GlyThrGlyAsnSerSerGlnValSerGlnAsnTyrrProIleValGlnAsnLeuGlnGly 140
QY 412 CAGATGCTGCACGAGGCATCAGCCCGCGACCTGAACGCTGGGTGAAGGTGATCAG 471
Db 141 GlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValValGlu 160
QY 472 GAGAAAGCCTTCAGCCCGAGGTGATCCCATGTTTACCGCCCTGAGCGAGGGCGCCACC 531
Db 161 GluLysAlaPheSerProGluValIleProMetPheSerAlaLeuSerGluGlyAlaThr 180
QY 532 CCCGAGACTGAACACAGATGTTGAACACCGTGGCGCGCACCGGCCCATGCAGATG 591
Db 181 ProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMet 200
```

```
QY 592 CTGAAGACACCATCAACGAGAGAGCCCGCGAGTGGAGCCGCTGCACCCCGTGCAGGCC 651
Db 201 LeuLysGluThrIleAsnGluGluAlaIleGluTrpAspArgLeuHisProValHisAla 220
QY 652 GGGCCCGTGGCCCCCGCGCAGATGCGCGAGACCCCGCGGACGACATCGCCGCGCCACC 711
Db 221 GlyProIleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThr 240
QY 712 AGCACCTGCAGAGACAGATCGCTGATGATGACCAACCCCGCTGCGCCGTGGCGAC 771
Db 241 SerThrLeuGlnGluGlnIleGlyTrpMetThrAsnAsnProIleProValGlyGlu 260
QY 772 ATCTACAAAGCGGTGATCATCTCTGGCGCTGAACAAGATCGTGGGATGTACAGCCCGTG 831
Db 261 IleTyrlsArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrsrProThr 280
QY 832 AGCATCTTGACATCCGCCAGGGCCCCAAGAGACCCCTCCGCGACTACGTGACCGCTTC 891
Db 281 SerIleLeuAspIleArgGlnGlyProLysGluProPheArgAspTyrlsValAspArgPhe 300
QY 892 TTCAAGACCCCTGGCGCGCGGAGCGAGCCACCCAGACGTAAGAACTGATGACCGAGACC 951
Db 301 TyrlsThrLeuArgAlaGluGlnAlaSerGlnGluValLysAsnTrpMetThrGluThr 320
QY 952 CTGCTGTGTGACAAAGCCCAACCCCGACTGCAAGACCATCTCGCGCGCTCTCGCGCCGCG 1011
Db 321 LeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyProAla 340
QY 1012 GCCACCTTGAGAGATGATGACCGCTGCGAGGGCGTGGCGCGCCCGCGCAAGGCC 1071
Db 341 AlaThrLeuGluGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLysAla 360
QY 1072 CGCGTCTGCGCGAGCGGATGAGCCAG--GCCAAGCGTGAACATCATGATGACAGAAG 1128
Db 361 ArgValLeuAlaGluAlaMetSerGlnValThrAsnSerAlaThrIleMetMetGlnArg 380
QY 1129 AGCACTTCAAGGGCCCCCGCGCAACGTCAAGTCTTCAACTGCGGCAAGAGGGCCAC 1188
Db 381 GlyAsnPheArgAsnGlnArgLysThrValLysCysPheAsnLysGlyLysValGlyHis 400
QY 1189 ATCGCCAAAGACTGCGCGCGCGCGCGCAAGAGGGCTGTGGAAGTGGCGCAAGAGGGC 1248
Db 401 IleAlaLysAsnLysCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGluGly 420
QY 1249 CACCAATGAAGGACTGCACCGGAGCGCCAGGCCAACTTCTGGGCAAGATCGGCCAGC 1308
Db 421 HisGlnMetLysAspCysAsnGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSer 440
QY 1309 CACAAGGGCGCGCGCGCAACTTCTGCAGAACCGGACGAGCGCGCGCCCGCCACCGTG 1368
Db 441 HisLysGlyArgProGlyAsnPheLeuGlnSerArgProGlu----- 454
QY 1369 CCCACCGCGCGCGCGCGGAGAGCTTCCGCTTC-----GAGGAGACCAACCCCGCCCCC 1422
Db 455 ProThrAlaProProGluGluSerPheArgPheGlyGluGluLysThrThrProSerGln 474
QY 1423 AAGCAGAGAGCCCAAGAGACCGCGAGCCCTACCGGAGACCCCTGACCGCCCTGCGGACCTG 1482
Db 475 LysGlnGluProIleAspLysGluLeuTyrr-----ProLeuAlaSerLeuArgSerLeu 492
QY 1483 TTGCGCAGCGCGCCCTGAGCCAG 1506
Db 493 PheGlyAsnAspProSerSerGln 500

RESULT 7
US-08-463-210-8
/ Sequence 8, Application US/08463210
/ Patent No. 6001977
/ GENERAL INFORMATION:
/ APPLICANT: CHANG, Nancy T.
/ APPLICANT: GALLO, Robert C.
/ APPLICANT: WONG-STAAI, Flossie
```

```
/ TITLE OF INVENTION: CLONING AND EXPRESSION OF HTLV-III DNA
/ NUMBER OF SEQUENCES: 11
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Morgan & Finnegan, L.L.P.
/ STREET: 345 Park Avenue
/ CITY: New York
/ STATE: New York
/ COUNTRY: USA
/ ZIP: 10154-0053
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent in Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/463,210
/ FILING DATE: 05-JUN-1995
/ CLASSIFICATION: 436
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 06/693,866
/ FILING DATE: 23-JAN-1985
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 06/659,339
/ FILING DATE: 10-OCT-1984
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Serunian, Leslie A.
/ REGISTRATION NUMBER: 35,353
/ REFERENCE/DOCKET NUMBER: 2026-4193US2
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 758-4800
/ TELEFAX: (212) 751-6849
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 512 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHEICAL: NO
/ ORIGINAL SOURCE:
/ ORGANISM: HTLV-III
/ FEATURE:
/ NAME/KEY: Protein
/ LOCATION: 1..512
/ OTHER INFORMATION: /note= "gag protein of HTLV-III"
/
US-08-463-210-8

Alignment Scores:
Pred. No.: 4.99e-128 Length: 512
Score: 2237.00 Matches: 422
Percent Similarity: 89.69% Conservative: 39
Best Local Similarity: 82.10% Mismatches: 39
Query Match: 78.77% Indels: 14
DB: 3 Gaps: 5

US-09-475-704A-4 (1-1509) x US-08-463-210-8 (1-512)
QY 1 ATGGGCGCGCGCGGAGCATCTGCGCGGCGAGAACTGGACAAGTGGAGAAATCCGC 60
Db 1 MetcylAlaArgAlaSerValLeuSerGlyGluLeuAspArgTrpGluLysIleArg 20
QY 61 CTGGCGCGCGCGCGGCAAGACATACTGTAAGCACTGTGTGGGCCAGCGCGAG 120
Db 21 LeuArgProGlyGlyLysLysLysLysLysLysLysLysLysLysLysLysLys 40
QY 121 CTGAGAGGCTTGGCCCTGAACCCCGGCTGTGAGACCGCGGAGGCTGCAAGAGATC 180
Db 41 LeuGluArgPheAlaValAsnProGlyLeuLeuGluThrSerGluGlyCysArgGlnIle 60
QY 181 ATGAAGCAGCTGCAGCCCGCGCTGAGACCGGACCGAGAGAGCTGGCAGCCTGTACAAC 240
Db 61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGluGluLeuAspGlySerLeuTyrAsn 80
```

```
QY 241 ACCGTGCCACCCCTGTAAGTGGTGCACGCCGGCATCGAGTCCCGGACACCAAGAGGCC 300
Db 81 ThrValAlaThrLeuTyrCysValHisGlnArgIleGluIleLysAspThrLysGluAla 100
QY 301 CTGGAACAAGATCGAGAGAGAGACAGAACAACTCCACAGACAGAACCCAGAGGCCAAGAG 360
Db 101 LeuAspLysIleGluGluGluGlnAsnLysSerLysLysValGlnGlnAlaAlaAla 120
QY 361 GCCGACGGC-----AAGTGAAGCCAGAACTAACCCCATCGTGACAACTTCAGAGGC 411
Db 121 AspThrGlyHisSerSerGlnValSerGlnAsnTyrProIleValGlnAsnIleGlnGly 140
QY 412 CAGATGCTGCACAGGCCCATCAGCCCCCGCACCTGAAACGCTGGGTGAAGTGAATCGAG 471
Db 141 GlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValValGlu 160
QY 472 GAGAAAGCCTTCAGCCCCCGAGGTGATCCCATGTTCACCCGCTGACGCGAGGCGCCACC 531
Db 161 GluLysAlaPheSerProGluValIlePrometPheSerAlaLeuSerGluGlyAlaThr 180
QY 532 CCCCAGGACCTGAAACAGATGTGAAACACCCGTGGGGCGGCCACAGCGCCCATGCAATG 591
Db 181 ProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMet 200
QY 592 CTGAAGACACCATCAACGAGAGAGCGCCGCGAGTGGGACCCGCTGCACCCCGTGACAGCC 651
Db 201 LeuLysGluThrIleAsnGluGluAlaAlaGluTrpAspArgValHisProValHisAla 220
QY 652 GGGCCCGTGGCCCCCGGCGAGATGCGCGACCCCGCGGAGAGCATCGCCGCGCCACC 711
Db 221 GlyProIleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThr 240
QY 712 AGCAACCTGCAGAGACAGATCGCGCTGGATGACCAACCAACCCCGCTGCGGGCGGAC 771
Db 241 SerThrLeuGlnGlnIleGlyTrpMetThrAsnAsnProIleProValGlyGlu 260
QY 772 ATCTACAAGCGGTGATCATCTGCGGCTGAAACAAGATCGTGCGGATGTACAGCCCGGTG 831
Db 261 IleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetLysSerProThr 280
QY 832 AGCATCTTGACATCCGCCAGGGCCCCAAGAGACCTTCGCGGACTACGTGACCGCTTC 891
Db 281 SerIleLeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhe 300
QY 892 TTCAAGACCCCTGGCGCGCGGAGAGGCGCACCCAGGACGTGAAGAACTGGATGACGAGACC 951
Db 301 TyrLysThrLeuArgAlaGluGlnAlaSerGlnGluValLysAsnTrpMetThrGluThr 320
QY 952 CTGCTGTGTGCAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGCGCCCGGC 1011
Db 321 LeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyProAla 340
QY 1012 GCCAACCTTGAGAGATGATGACCGCTGCGAGGGCGGTGGGGGGCCCCGGGCAACAAGGCC 1071
Db 341 AlaThrLeuGluGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLysAla 360
QY 1072 CGCGTGTGCGCGGAGGCGATGAGCCAG--GCCAAGCGGTGAACATCATGATGCAGAAAG 1128
Db 361 ArgValLeuAlaGluAlaMetSerGlnValThrAsnThrAlaThrIleMetMetGlnArg 380
QY 1129 AGCAACTTCAAGGGCCCCCGCGCGCAACGTCAAGTCTTCAACTGCGGCAAGAGGGCCAC 1188
Db 381 GlyAsnPheArgAsnGlnArgLysMetValLysCysPheAsnCysGlyLysGluGlyHis 400
QY 1189 ATGCCCAAGAACTGCGCGCGCGCGCGCAAGAGGGCTGTGAAAGTGCGGCAAGAGGGGC 1248
Db 401 ThrAlaArgAsnCysArgAlaProArgLysLysGlyCysThrLysCysGlyLysGluGly 420
QY 1249 CACCAGATGAAGACTGCACGAGCGCGGCAAGCTTCTGCGCAAGATGTGCGCCAGC 1308
Db 421 HisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSer 440
QY 1309 CACAAGGGCGCGCGCGCAACTTCTGCAAGAACCGAGCGAGCCGCGCGCCACCGCTG 1368
```

```
Db      441 TyrLysGIYArGProGIYAsnPheLeuGlnSerArGProGIUProThrAlaProPhe 460
QY      1369 -----CCCAACGGCCCCCGCCGAGAGCTTCCGC-----TTCGAG 1404
Db      461 LeuGlnSerArGProGIUProThrAlaProGIUInSerPheArGSerGIYValGIU 480
QY      1405 GAGACCAACCCCGCCCAAGAGAGCCCAAGACCGGAGCCCTACCGGAGCCCTG 1464
Db      481 ThrThrThrProGIUInSerGIUInGIUProIleAspLysGIULeuTyr-----ProLeu 498
QY      1465 ACCGCCCTGCGAGCCTGTTGCGAGCGGCCCTGAGCCAG 1506
Db      499 ThrSerLeuArGSerLeuPheGIYAsnAspProSerSerGIU 512

RESULT 8
US-09-124-900-2
; Sequence 2, Application US/09124900
; Patent No. 6268484
; GENERAL INFORMATION:
; APPLICANT: KATINGER, Hermann
; APPLICANT: BUCHACHER, Andrea
; APPLICANT: ERNST, Wolfgang
; APPLICANT: BALLAUN, Claudia
; APPLICANT: PURTSCHER, Martin
; APPLICANT: TRKOLA, Alexandra
; APPLICANT: PREDL, Renate
; APPLICANT: SCHMATZ, Christine
; APPLICANT: KLIMA, Annelies
; APPLICANT: STEINDL, Franz
; APPLICANT: MUSTER, Thomas
; TITLE OF INVENTION: HIV-Vaccines
; FILE REFERENCE: 1939-112P
; CURRENT APPLICATION NUMBER: US/09/124, 900
; CURRENT FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: PCT/EP95/01481
; PRIOR FILING DATE: 1995-04-19
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-124-900-2

Alignment Scores:
Pred. No.:      4.99e-128      Length:      512
Score:          2237.00      Matches:      422
Percent Similarity: 89.69%      Conservative: 39
Best Local Similarity: 82.10%      Mismatches: 39
Query Match:    78.77%      Indels:      14
DB:             3          Gaps:      5

US-09-475-704A-4 (1-1509) x US-09-124-900-2 (1-512)
QY      1 ATGGGCGCGCGCGCCAGCATCTGCGCGCGAGAGAGCTGACAAAGTGGAGAAGATCCGC 60
Db      1 MetGIYAlaArGAlaSerValLeuSerGIYGIULeuAspArGTpGIUlysIleArG 20
QY      61 CTGGCGCGCGCGCGGAGAAAGCACTGATGCTGAAGCACTGTGTGGGCCAGCGCGAG 120
Db      21 LeuArGProGIYGIYsLysLysTyrLysLeuLysHisIleValTrpAlaSerArGIU 40
QY      121 CTGAGGGCTTGCCTGAAACCCCGGCTGTGAGAGCCGCGAGGGCTGCAAGCAGATC 180
Db      41 LeuGIUArGpHeAlaValAsnProGIYLeuLeuGIUThrSerGIUInGIYCsArGIU 60
QY      181 ATGAAGAGCTGACGCCCGCTGACAGCCGACCGAGAGCTGCGAGCCTGTACAAC 240
Db      61 LeuGIYGIUInLeuGIUProSerLeuGIUInThrGIYSerGIUInLeuArGSerLeuTyrAsn 80
QY      241 ACCGTGGCCACCCCTGACTGCGTGCAGCGCCGCATCGAGGTCCGCGACACCAAGAGGCC 300
|||||
```

```
Db      81 ThrValAlaThrLeuTyrCysValHisGIUArGIleGIUlleLysAspThrLysGIUAla 100
QY      301 CTGGAACAAGATCGAGGAGGAGCAGACAACAGTCCGACGAGAGACCCAGAGGCCAAGAG 360
Db      101 LeuAspLysIleGIUInGIUInGIUInAsnLysSerLysLysAlaGIUInGIUAlaAla 120
QY      361 GCCGACGGC-----AAGGTAGCCAGAACTAACCCCATCGTGCAAGAACTGCAAGGGC 411
Db      121 AspThrGIYHisSerSerGIUInValSerGIUInAsnTyrProIleValGIUInAsnIleGIU 140
QY      412 CAGATGGTGACACGAGCCATCAGCCCCCGACACCTGAACGCGCTGGGTGAAGGTATGAG 471
Db      141 GIUInValHisGIUInAlaIleSerProArGIUInLeuAsnAlaTrpValLysValGIU 160
QY      472 GAGAGGCCCTTACGCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGGCGCCACC 531
Db      161 GIUInLysAlaPheSerProGIUInValIleProMetPheSerAlaLeuSerGIUInGIUAlaThr 180
QY      532 CCCGAGGACCTGAACACGATGTTGAACACCGTGGGCGGCCACGAGCCGCGCATGAGATG 591
Db      181 ProGIUInAspLeuAsnThrMetLeuAsnThrValGIYGIUInHisGIUInAlaIleMetGIU 200
QY      592 CTGAAGACACCACTCAACGAGAGGCGCGCGAGTGGAGCCGCTGCAACCCGTCAGAGCC 651
Db      201 LeuLysGIUThrIleAsnGIUInAlaIleGIUInTrpAspArGIUInHisProValHisAla 220
QY      652 GGGCCCGTGGCCCCCGCCAGATGCGCGAGACCCCGCGGCGAGCGACATGCGCGCGCCACC 711
Db      221 GIYProIleAlaProGIYGIUInMetArGIUProArGIYSerAspIleAlaGIYThrThr 240
QY      712 AGCACCTTGACAGAGCAGATGCGCTGGATGACCAACACCCCGCTGCGCGTGGCGGAC 771
Db      241 SerThrLeuGIUInGIUInIleGIYTrpMetThrAsnAsnProIleProValGIYGIU 260
QY      772 ATCTCAAGCGGTGATCATCTCTGGGCTGAAACAAGATCGTGGCATGTACAGCCCGTG 831
Db      261 IleTyrLysArGTpIleIleLeuGIYLeuAsnLysIleValArGIUInMetTyrSerProThr 280
QY      832 AGCATCCTGACATCCGCCAGGCGCCCAAGAGACCCCTCCGCGACTACGTGACCGCTTC 891
Db      281 SerIleLeuAspIleArGIUInGIYProLysGIUProPheArGIUInTrpValAspArGIU 300
QY      892 TTCAAGACCCCTGCGCGCGCGAGAGCCACCCAGAGCCTGAAGAACTGATGACCGAGACC 951
Db      301 TyrLysThrLeuArGIUInGIUInAlaIleSerGIUInGIUValLysAsnTrpMetThrGIUThr 320
QY      952 CTGCTGTGTCAGAACGCCCAACCCCGACTGCAAGAACCATCTGCGCGCTCTCGGCCCGGC 1011
Db      321 LeuLeuValGIUInAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGIYProAla 340
QY      1012 GCCACCTTGAGAGCATGATGACCGCTGCCAGGCGGTGGCGGCCCGCCACAGAGGCC 1071
Db      341 AlaThrLeuGIUInGIUInMetMetThrAlaCysGIUInGIYValGIYGIYProGIYHisLysAla 360
QY      1072 CGCGTGTGCGCGAGGCGGATGAGCCAG--GCCAACAGCGTGAACATCATGATGACAGAAG 1128
Db      361 ArGIUInValLeuAlaGIUInAlaMetSerGIUInValThrAsnThrAlaThrIleMetMetGIUInArG 380
QY      1129 AGCAACTTCAAGGGCCCCCGCGCGCAACGTCAAGTGTCTTCACTGCGGCAAGAGAGGCCAC 1188
Db      381 GIYAsnPheArGIUInGIUInArGIUInLysMetValLysCysPheAsnGIYGIYLysGIUInHis 400
QY      1189 ATCGCCAAGAACTGCGCGCGCGCGCGCAAGAGGGCTGTGAAAGTGGGCAAGATCTGGCCCCAGC 1248
Db      401 ThrAlaArGIUInCysArGIUInAlaProArGIUInLysGIYCysTrpLysCysGIYLysGIUIn 420
QY      1249 CACCAATGAAGAGCTGACCGGAGCGCCAGGCCAACTTCTCTGGGCAAGATCTGGCCCCAGC 1308
Db      421 HisGIUInMetLysAspCysThrGIUInArGIUInAlaAsnPheLeuGIYLysIleTrpProSer 440
QY      1309 CACAAGGGCGCGCGCGCAACTTCTTGCAGAACCGCAGAGCGCGCGCGCCCAACCGGTG 1368
Db      441 TyrLysGIYArGProGIYAsnPheLeuGIUInSerArGIUProGIUProThrAlaProPhe 460
|||||
```

QY 1369 -----CCCCCCCCCCCCCGCGAGAGCTTCCGC-----TTCGAG 1404
Db 461 LeuGlnSerArgProGluProThrAlaProGluGlnSerPheArgSerGlyValGlu 480
QY 1405 GAGACCAACCCCGCCCAAGAGAGCCCAAGAGCGAGCCCTAACCGCGAGCCCTG 1464
Db 481 ThrThrThrProGlnGlnGlnProIleAspGlyGluLeuTyr-----ProLeu 498
QY 1465 ACCGCCCTGCGAGCTGTTCGAGCGGCGCCCTGAGCCAG 1506
Db 499 ThrSerLeuArgSerLeuPheGlyAsnAspProSerSerGln 512
RESULT 9
US-08-463-028-8
; Sequence 8, Application US/08463028
; Patent No. 6610476
; GENERAL INFORMATION:
; APPLICANT: CHANG, Nancy T.
; APPLICANT: GALLO, Robert C.
; APPLICANT: WONG-STAL, Flossie
; TITLE OF INVENTION: CLONING AND EXPRESSION OF HTLV-III DNA
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,028
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/693,866
; FILING DATE: 23-JAN-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/659,339
; FILING DATE: 10-OCT-1984
; ATTORNEY/AGENT INFORMATION:
; NAME: Serunian, Leslie A.
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-4193US3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 512 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: HTLV-III
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..512
; OTHER INFORMATION: /note= "gag protein of HTLV-III"
US-08-463-028-8

Alignment Scores:
Pred. No.: 4.99e-128 Length: 512
Score: 2237.00 Matches: 422
Percent Similarity: 89.69% Conservative: 39
Best Local Similarity: 82.10% Mismatches: 39

Query Match: 78.77% Indels: 14
DB: 4 Gaps: 5
US-09-475-704A-4 (1-1509) x US-08-463-028-8 (1-512)
QY 1 ATGGGCGCCGCGCCAGCATCTGCGCGCGGAGAGCTGGACAAGTGGAGATCCGC 60
Db 1 MetGlyAlaArgAlaSerValLeuSerGlyGlyLeuAspArgTrpGluValArg 20
QY 61 CTGGCGCCCGCGCGCAAGAGCACTATGCTGAAGCACCTGTGTGGGCCAGCCGCGAG 120
Db 21 LeuArgProGlyGlyLysLysLysTyrLysLeuLysHisIleValTrpAlaSerArgGlu 40
QY 121 CTGGAGGGCTTCGCCCCGTGAACCCCGCGCTGTGAGAGACCGCGAGGGCTGACAGATC 180
Db 41 LeuGluArgPheAlaValAlaAsnProGlyLeuLeuGluThrSerGluGlyCysArgGlnIle 60
QY 181 ATGAAGAGCTGAGCCCGCGCTGCAGACCGGCAACCGAGAGCTGCAGCTGTACAAC 240
Db 61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGluGluLeuArgSerLeuTyrAsn 80
QY 241 ACCGTGGCCACCTGTACTGCTGCAGCGCGGCATCGAGGTCCGCAACCAAGAGGCC 300
Db 81 ThrValAlaThrLeuTyrCysValHisGlnArgGlyIleGluLysAspThrLysGluAla 100
QY 301 CTGGAACAAGATGAGAGAGAGAGAGCAACAAGTCCCGACAGAGAACCCAGAGCCCAAGAG 360
Db 101 LeuAspLysIleGluGluGluGluGlnAsnLysSerLysLysValAlaGlnAlaAla 120
QY 361 GCCGACGCGC-----AAGGTAGCCAGAACTACCCCATGCTGCAGAACTGCAGAGGC 411
Db 121 AspThrGlyHisSerSerGlnValSerGlnAsnTyrProIleValGlnAsnIleGlnGly 140
QY 412 CAGATGCTGACACGAGCCATGACGCCCCCGCACCTGAACGCTGGGTGAAGGTGATCGAG 471
Db 141 GlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValGlu 160
QY 472 GAGAAAGCCTTCAGCCCCGAGGTGATCCCATGTTCAACCGCCTGAGCGAGGCGCCACC 531
Db 161 GluLysAlaPheSerProGluValIleProMetPheSerAlaLeuSerGluGlyAlaThr 180
QY 532 CCCCAGGACCTGAAACAGATGTTGAACACCGTGGCGGCGCAACGCGCCATGACAGATG 591
Db 181 ProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaIleMetGlnMet 200
QY 592 CTGAAGAGACCATCAACGAGAGAGGCGCGCGAGTGGAGACCGCCTGCACCCCGTCAGGCC 651
Db 201 LeuLysGluThrIleAsnGluGluAlaIleGluTrpAspArgValHisProValHisAla 220
QY 652 GGGCCCCGTGGCCCCCGCCAGATGCGCGGACCCCGCGGAGAGCATCGCGCGCGCCACC 711
Db 221 GlyProIleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThr 240
QY 712 AGCAACCTGCAGAGACAGATCGCCTGGATGACCAAGCAACCCCGCGTGGCGGCGAC 771
Db 241 SerThrLeuGlnGluGlnIleGlyTrpMetThrAsnAsnProIleProValGlyGlu 260
QY 772 ATCTACAAGCGGTGATCATCTCGGGCGCTGAACAAGATCGTGGGATGTACAGCCCGCGT 831
Db 261 IleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProThr 280
QY 832 AGCATCTTGACATCCGCCAGGAGCCCAAGAGAGCCCTTCCGCGACTACGTGACCGCTTC 891
Db 281 SerIleLeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhe 300
QY 892 TTCAGAAGCCCTGCGCGCGCGAGAGAGGCCAACCAAGAGAGCTGTAAGAACTGATGACCGAGACC 951
Db 301 TyrLysThrLeuArgAlaGluGlnAlaSerGlnGluValLysAsnTrpMetThrGluThr 320
QY 952 CTGCTGTGTGACAAGCGCAACCGCCGACTGCAAGAACCATCTCGCGCGCTCTGGCGCCCGGC 1011
Db 321 LeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyProAla 340

```
QY 1012 GCCACCTTGAGAGATGATGACCGCTGCGAGGGCGTGGCGCCCGCCCAAGGCC 1071
Db 341 A1atThrLeuGluGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLysAla 360
QY 1072 CGCGTGTGCGCGAGCGCATGAGCCAG---GCCAACAGCGTGAACATCATGATGACAGAG 1128
Db 361 ArgValLeuAlaGluAlaMetSerGlnValThrAsnThrAlaThrIleMetMetGlnArg 380
QY 1129 AGCAACTTCAAGGCGCCCGCGCAACGTCAGTCTTCACTGCGGCAAGAGGCGCAC 1188
Db 381 G1yAsnPheArgAsnGlnIleArgLysMetValLysCysPheAsnCysGlyLysGluGlyHis 400
QY 1189 ATGCGCAAGAACTGCGCGCGCCCGCGCAAGAGGCTGCTGAAGTGGCGCAAGAGGCGC 1248
Db 401 ThrAlaArgAsnCysArgAlaProArgLysLysGlyCysThrPlysCysGlyLysGluGly 420
QY 1249 CACCAGATGAGAGACTGACACGAGCGCGCAACTTCTGCGCAAGATCTGGCCAGC 1308
Db 421 HisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSer 440
QY 1309 CACAAGGCGCGCGCGCAACTTCTGCAAGACCGCAGAGCGCGCGCCCGCCACCGTG 1368
Db 441 TyrLysGlyArgProGlyAsnPheLeuGlnSerArgProGluProThrAlaProProPhe 460
QY 1369 -----CCCAACGCGCGCGCGCGCGAGAGCTTCCGC-----TTGCAG 1404
Db 461 LeuGlnSerArgProGluProThrAlaProProGluGlnSerPheArgSerGlyValGlu 480
QY 1405 GAGACCAACCCCGCGCGCGCAAGAGAGCCCAAGAGACCGCGAGCCCTACCGGAGCCCTG 1464
Db 481 ThrThrThrProProGlnLysGlnGluProIleAspLysGluLeuTyr-----ProLeu 498
QY 1465 ACCGCGCTGCGCAGCCTGTTCGGCAGCGCGCCCTGAGCCAG 1506
Db 499 ThrSerLeuArgSerLeuPheGlyAsnAspProSerSerGln 512

RESULT 10
US-08-816-155B-45
; Sequence 45, Application US/08816155B
; Patent No. 5990091
; GENERAL INFORMATION:
; APPLICANT: TARTAGLIA, JAMES
; APPLICANT: COX, WILLIAM I.
; APPLICANT: GETTIG, RUSSELL R.
; APPLICANT: MARTINEZ, HECTOR
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: PINCUS, STEVEN E.
; TITLE OF INVENTION: VECTORS HAVING ENHANCED EXPRESSION, AND
; TITLE OF INVENTION: METHODS OF MAKING AND USES THEREOF
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP
; STREET: 745 FIFTH AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10151
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/816,155B
; FILING DATE: 12-MAR-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: KOMALSKI, THOMAS J.
; REGISTRATION NUMBER: 32,147
; REFERENCE/DOCKET NUMBER: 454310-2990
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-588-0800
; TELEFAX: 212-588-0500
```

```
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 500 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-816-155B-45

Alignment Scores:
Pred. No.: 3.51e-127 Length: 500
Score: 2223.00 Matches: 420
Percent Similarity: 90.16% Conservative: 38
Best Local Similarity: 82.68% Mismatches: 36
Query Match: 78.27% Indels: 14
DB: 2 Gaps: 5

US-09-475-704A-4 (1-1509) x US-08-816-155B-45 (1-500)
QY 1 ATGGGCGCGCGCGCGCATCTGCGCGCGAGAACTGGAACAAGTGGAGAAATCCGC 60
Db 1 MetGlyAlaArgAlaSerValLeuSerGlyGlyGluLeuAspArgTrpGluLysIleArg 20
QY 61 CTGCGCGCGCGCGCGCAAGACACTATCATGTGAAGCACTGTGTGGCCAGCGCGAG 120
Db 21 LeuArgProGlyGlyLysLysLysTyrLysLeuLysHisIleValIlePheAlaSerArgGlu 40
QY 121 CTGAGGGCTTTCGCTGAACCCCGCGCTGTGAGACCGCGAGGGCTGCAAGCAGATC 180
Db 41 LeuGluArgPheAlaValAlaAsnProGlyLeuLeuGluThrSerGluGlyCysArgGlnIle 60
QY 181 ATGAAGCAGCTGACCGCGCGCGCTGAGACCGCGCAGAGAGAGTGGCGAGCCTGTACAAC 240
Db 61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGluGluLeuArgSerLeuTyrAsn 80
QY 241 ACCGTGGCACCCCTGTACTGTGCTGACCGCGCATTCGAGGTCCGACCAAGAGAGCC 300
Db 81 ThrValAlaThrLeuTyrCysValHisGlnArgIleGluIleLysAspThrLysGlnAla 100
QY 301 CTGCAAGATCGAGAGAGAGAGAGAGCAAGTCCGACAGAGAGAGAGAGAGAGAGAG 360
Db 101 LeuAspLysIleGluGluGluGlnAsnLysSerLysLysAlaGlnGlnAlaAla 120
QY 361 GCCGAGCGC-----AAGGTGAGCCAGAACTACCCCATCGTCAGAACTGCGAGGC 411
Db 121 AspThrGlyHisSerAsnGlnValSerGlnAsnTyrProIleValGlnAsnIleGlnGly 140
QY 412 CAGATGTTGACACGAGCCATCAGCCCGCGACCCCTGAAGCCTGGGTGAAGTGAATCAG 471
Db 141 GlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaIleTrpValLysValGlu 160
QY 472 GAGAGGCCTTACGCGCGCGAGGTGATCCCATGTTCAACGCGCCTGAGCGAGCGCAC 531
Db 161 GluLysAlaPheSerProGluValIlePrometPheSerAlaLeuSerGluGlyAlaThr 180
QY 532 CCCGAGACCTGAACAGATGTTGAACACCGTGGCGGCCACCGCGCCCATGAGATG 591
Db 181 ProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMet 200
QY 592 CTGAAGACACCATCAACGAGAGGCGCGCGAGTGGAGACCGCCTGACCCCGTGAGGCC 651
Db 201 LeuLysGluThrIleAsnGluGlnAlaIleGluTrpAspArgValHisProValHisAla 220
QY 652 GCGCGCGTGGCGCGCGCGCAGATGCGCGACCCCGCGGCGAGCGACATCGCGCGCAC 711
Db 221 GlyProIleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThr 240
QY 712 AGCACTGACAGAGACAGATCGCGCTGATGACGACCAACCCCGCGTGGCGCGAC 771
Db 241 SerThrLeuGlnGlnIleGlyTyrMetThrAsnAsnProProIleProValGlyGlu 260
QY 772 ATCTACAAGCGGTGATCATCTCGGCGCTGAACAAGATCGTGGCATGTACAGCCCGCTG 831
```


Db 261 IleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProThr 280
 QY 832 AGCATCCTGGACATCCGCGCAGAGGCCCCGAGAGAGCCCTTCCGCGACTACGTGAGACCGCTTC 891
 Db 281 SerIleuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhe 300
 QY 892 TTCAAGACCCTGCGCGCGAGCAGGCCACCCAGAGACGTGAAGAACTGGATGACCGAGACC 951
 Db 301 TyrLysThrLeuArgAlaGlnAlaIleSerGlnGluValLysAsnTrpMetThrGluThr 320
 QY 952 CTGCTGGTGCAGAACCGCAACCCCGACTGCAAGACCACTCGCGCTCTCGGCCCGCCG 1011
 Db 321 LeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyProAla 340
 QY 1012 GCCACCTGGAGAGATGATGACCGCTGCGCAGGCGCTGGCGGCCCGCCGACAGGCC 1071
 Db 341 AlaThrLeuGlnGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLysAla 360
 QY 1072 CGCGTCTGGCCGAGCGATGAGCCAG---GCCAAGACGCTGAACATCATGATGACAGAG 1128
 Db 361 ArgValLeuAlaGlnAlaMetSerGlnValThrAsnSerAlaThrIleMetMetGlnArg 380
 QY 1129 AGCAACTTCAAGGCGCCCGCGCAACGTCAAGTCTTCACTGCGCGCAAGAGGCGCAC 1188
 Db 381 GlyAsnPheArgAsnGlnArgLysIleValLysCysPheAsnCysGlyLysGluGlyHis 400
 QY 1189 ATCGCCAGAACTGCGCGCGCCCGCGCAAGAGGCTGCTGGAAGTGGCGCAAGAGGCGC 1248
 Db 401 ThrAlaArgAsnCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGluGly 420
 QY 1249 CACCAAGTGAAGACTGACCGAGCGCGCACTTCTGCGCAAGATCTGCGCCAGC 1308
 Db 421 HisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSer 440
 QY 1309 CACAAGGCGCGCGCGCAACTTCTGCAAGACCGCAGCGAGCCCGCGCCCGCCACCGCTG 1368
 Db 441 TyrLysGlyArgProGlyAsnPheLeuGlnSerArgProGlu----- 454
 QY 1369 CCCACCGCGCGCGCGCGAGAGCTTCCGC-----TTCCAGAGAGACCAACCCCGCGCCC 1422
 Db 455 ProThrAlaProProGluGlnSerPheArgSerGlyValGluThrThrProProGln 474
 QY 1423 AAGCAGAGAGCCCAAGAGCCGCGAGCCCTTACCGCGAGCCCTGACCGCGCTGCGAGCCTG 1482
 Db 475 LysGlnGlnProLysAspLysGlnLeuTyr-----ProLeuThrSerLeuArgSerLeu 492
 QY 1483 TTCGCGAGCGCGCGCGCTGAGCCAG 1506
 Db 493 PheGlyAsnAspProSerSerGln 500

RESULT 11

US-09-079-587-45
 ; Sequence 45, Application US/09079587
 ; Patent No. 6130066
 ; GENERAL INFORMATION:
 ; APPLICANT: TARTAGLIA, JAMES
 ; APPLICANT: COX, WILLIAM I.
 ; APPLICANT: GETTING, RUSSELL R.
 ; APPLICANT: MARTINEZ, HECTOR
 ; APPLICANT: PAOLETTI, ENZO
 ; APPLICANT: PINCUS, STEVEN E.
 ; TITLE OF INVENTION: VECTORS HAVING ENHANCED EXPRESSION, AND
 ; TITLE OF INVENTION: METHODS OF MAKING AND USES THEREOF
 ; NUMBER OF SEQUENCES: 48
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP
 ; STREET: 745 FIFTH AVENUE
 ; CITY: NEW YORK
 ; STATE: NEW YORK
 ; COUNTRY: USA
 ; ZIP: 10151
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/079,587
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/816,155
 FILING DATE: 12-MAR-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: KOWALSKI, THOMAS J.
 REGISTRATION NUMBER: 32,147
 REFERENCE/DOCKET NUMBER: 454310-2990
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-588-0800
 TELEFAX: 212-588-0500
 INFORMATION FOR SEQ ID NO: 45:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 500 amino acids
 TYPE: amino acid
 STRANDEDNESS: n/a
 TOPOLOGY: linear
 MOLECULE TYPE: amino acid
 US-09-079-587-45

Alignment Scores:
 Pred. No.: 3,51e-127 Length: 500
 Score: 2223.00 Matches: 420
 Percent Similarity: 90.16% Conservative: 38
 Best Local Similarity: 82.68% Mismatches: 36
 Query Match: 78.27% Indels: 14
 DB: 3 Gaps: 5

US-09-475-704A-4 (1-1509) x US-09-079-587-45 (1-500)

QY 1 ATGGGCGCGCGCGCGCAGCATCTGCGCGCGCGAGAACTGGACAAGTGGAGAAATCCGC 60
 Db 1 MetGlyAlaArgAlaSerValLeuSerGlyGlyGluLeuAspArgTrpGluLysIleArg 20
 QY 61 CTGGCGCGCGCGCGCAAGAGCACTACATGCTGAAGCAACCTGTGTGGCCAGCCGCGAG 120
 Db 21 LeuArgProGlyGlyLysLysLysTyrLysLeuLysHisIleValTrpAlaSerArgGlu 40
 QY 121 CTGGAGGGCTTGGCCCTGAACCCCGCGCTGCTGGAAGACCGCGGAGGGCTGCAAGCATC 180
 Db 41 LeuGluArgPheAlaValAsnProGlyLeuLeuGluThrSerGluGlyCysArgGlnIle 60
 QY 181 ATGAAGCAGCTGAGACCCCGCGCTGCAAGACCGGACCGAGAGCTGCGAGCCTGTACAAC 240
 Db 61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGluGluLeuArgSerLeuTyrAsn 80
 QY 241 ACCGTGGCCACCTGTACTGCTGTCACGCGCGCATGAGGTCCGACACCAAGAGAGGCC 300
 Db 81 ThrValAlaThrLeuTyrCysValHisGlnArgIleGluIleLysAspThrLysGluVal 100
 QY 301 CTGGAACAAGATCAGAGAGAGAGCAACAAGTCCAGCAGAGAACCAGAGCGCAAGAGAG 360
 Db 101 LeuAspLysIleGluGluGluGlnAsnLysSerLysLysValAlaGlnAlaAla 120
 QY 361 GCCGACGGC-----AAGGTGAGCCAGAACTACCCCATGCTGCAGAACTGCAGGGC 411
 Db 121 AspThrGlyHisSerAsnGlnValSerGlnAsnTyrProIleValGlnAsnIleGlnGly 140
 QY 412 CAGATGCTGACCAAGCCATCAGCCCGCGACCTGAACGCTGGTGAAGGTGATCGAG 471
 Db 141 GlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValGlu 160
 QY 472 GAGAAAGCCTTACAGCCCGAGGTGATCCCATGTTCAACCGCCTGAGCGAGCGCCACC 531
 Db 161 GlutLysAlaPheSerProGluValIlePrometPheSerAlaLeuSerGluGlyAlaThr 180
 QY 532 CCCCAGGACCTGAACAGATGTTGAACACCGTGGCGGCGCACAGCGCCGCGCATGAGATG 591

```
Db      181 ProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaIleMetGlnMet 200
QY      592 CTGAAGACACCATCAACGAGAGCGCCGAGTGGACCGCCTGCACCCCGTGCAGGCC 651
Db      201 LeuYsGluThrIleAsnGlnGlnAlaIleGlnTrpAspArgValHisProValHisAla 220
QY      652 GGGCCCGTGGCCCCCGCCAGATGCGCGCACCCCGCGCAGCAGACATCGCCGCGCAC 711
Db      221 GlyProIleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThr 240
QY      712 AGCACCTTGACGAGCAGATCGCCTGGATGACCAACCCCCCGTGCCTGGCGGAC 771
Db      241 SerThrLeuGlnGlnGlnIleGlyTrpMetThrAsnAsnProProIleProValGlyGlu 260
QY      772 ATCTACAAGCGGTGATCATCTGGCGCTGAACAAGATCGTGGATGTACAGCCCGTG 831
Db      261 IleTyrIysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProThr 280
QY      832 AGCATCTTGACATCCGCGCAGAGCGCCCAAGAGAGCCCTTCGCGACTACGTGACCGCTTC 891
Db      281 SerIleLeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhe 300
QY      892 TTCAAGACCTTGCGCGCGCAGACGCGCACCCAGACGTCGAAGATCGATGACCGAGACC 951
Db      301 TyrLysThrLeuArgAlaGlnGlnAlaSerGlnGlyValLysAsnTrpMetThrGluThr 320
QY      952 CTGCTGTGCAGAACGCCAACCCCGCATGCAAGCATCTCGCGCTCTCGCGCCCGGC 1011
Db      321 LeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyProAla 340
QY      1012 GCCACCTTGAGAGATGATGACCGCGCTGCCAGCGCTGGCGCGCCCGCCCAAGAGGCC 1071
Db      341 AlaThrLeuGlnGlnMetMetThrAlaCysGlnGlyValGlyProGlyHisLysAla 360
QY      1072 CGCGTCTGGCCGAGCGCATGAGCCAG--GCCAAGACCGTGAACATCATGATGCAGAAG 1128
Db      361 ArgValLeuAlaGlnAlaMetSerGlnValThrAsnSerAlaThrIleMetMetGlnArg 380
QY      1129 AGCAACTTCAGAGGCGCGCGCGCAACGTCAAGTCTTCAACTGCGGCAAGAGGGCCAC 1188
Db      381 GlyAsnPheArgAsnGlnArgLysIleValLysCysPheAsnCysGlyLysGlnGlyHis 400
QY      1189 ATCCGCCAAGAACTGCCGCGCGCGCGCGCAAGAGGGCTGTGAAGTGGCGGCAAGAGGGCC 1248
Db      401 ThrAlaArgAsnCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGlnGly 420
QY      1249 CACCAATGAAGGACTGCACCGAGCGCGCAAGGCAACTTCTGGGCAAGATTTGGCCAGC 1308
Db      421 HisGlnMetLysAspCysThrGlnArgGlnAlaAsnPheLeuGlyLysIleTrpProSer 440
QY      1309 CACAAGGGCGCGCGCGCAACTTCTGTGCAGAACCGCAGCGCGCGCGCCCAACCGTG 1368
Db      441 TyrLysGlyArgProGlyAsnPheLeuGlnSerArgProGlu----- 454
QY      1369 CCCACGCGCGCGCGCGCGAGAGCTTCGCG-----TTCGAGGAGACCAACCCCGCGCCC 1422
Db      455 ProThrAlaProProGlnGlnLysSerPheArgSerGlyValGlnThrThrProProGln 474
QY      1423 AAGCAGAGCGCCCAAGAGCGCGAGCGCTTACCGGAGCCCTGCAGCCCTGCGCAGCCTG 1482
Db      475 LysGlnGlnLysProIleAspLysGlnLeuTyr-----ProLeuThrSerLeuArgSerLeu 492
QY      1483 TTCGCGACGCGCGCGCTGAGCCAG 1506
Db      493 PheGlyAsnAspProSerSerGln 500
```

```
; APPLICANT: Youil, Rima
; APPLICANT: Bett, Andrew J.
; APPLICANT: Chen, Ling
; APPLICANT: Kaslow, David C.
; APPLICANT: Shiver, John W.
; APPLICANT: Toner, Timothy J.
; APPLICANT: Casimiro, Danilo R.
; TITLE OF INVENTION: ENHANCED FIRST GENERATION ADENOVIRUS
; TITLE OF INVENTION: VACCINES EXPRESSING CODON OPTIMIZED HIV1-GAG, POL, NEF AND
; TITLE OF INVENTION: MODIFICATIONS
; FILE REFERENCE: 20747Y
; CURRENT APPLICATION NUMBER: US/09/952,060
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: PCT/US01/28861
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/317,814
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/279,056
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/233,180
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: tPA-gag fusion open reading frame
US-09-952-060-33
```

```
Alignment Scores:
Pred. No.:      2,29e-119      Length:      493
Score:          2094.00        Matches:      396
Percent Similarity: 90.53%      Conservative: 34
Best Local Similarity: 83.37%      Mismatches:  31
Query Match:      73.73%        Indels:      14
DB:              4            Gaps:          5
```

```
US-09-475-704A-4 (1-1509) x US-09-952-060-33 (1-493)
QY      100 CTGCTGTGGGCCACGCGCGAGCTGAGAGGGCTTCGCTGAACCCCGGCTGTGAGAGCC 159
Db      27 ILeValTrpAlaSerArgGlnLeuGlnArgPheAlaValAsnProGlyLeuLeuGluThr 46
QY      160 GCGGAGGCTGCAAGCAGATCATGAAGCAGCTGCAGCCCGCCTGCAGACCGGACCGAG 219
Db      47 SerGlnGlyCysArgGlnIleLeuGlnLeuGlnProSerLeuGlnThrGlySerGlu 66
QY      220 GAGCTGCGACGCTGTACAACCGTGGCGCACCTGTACTGCGTGCACGCGCGCATCGAG 279
Db      67 GluLeuArgSerLeuTyrAsnThrValAlaThrLeuTyrCysValHisGlnLysIleAsp 86
QY      280 GTCCGCGACACCAAGAGAGCGCTTGACAAGATCGAGAGCAGACGAACAAGTCCAGCAG 339
Db      87 ValLysAspThrLysGlnAlaLeuGlnLysIleGlnGlnGlnAsnLysSerLysLys 106
QY      340 AAGACCCAGACGCGCCAAAGAGGCGCGAGC-----AAGGTGAGCCAGAACTAACCCC 390
Db      107 LysAlaGlnGlnAlaAlaAlaGlyThrGlyAsnSerSerGlnValSerGlnAsnTyrPro 126
QY      391 ATCGTGAGAACTTGACAGGCGCAGATGTGTGACCAAGCATCAGCCCGCGACCTGAAC 450
Db      127 ILeValGlnAsnLeuGlnGlyGlnMetValHisGlnAlaIleSerProArgThrLeuAsn 146
QY      451 GCCTGGGTGAAGTGTATCGAGAGAGAGGCTTCAGCCCGAGGTGATCCCATGTTCAAC 510
Db      147 AlaTrpValLysValValGlnGlnLysAlaPheSerProGlnValIleProMetPheSer 166
QY      511 GCCCTGAGCGAGGGCGCCACCCCGACGACTGAACACGATGTGAACACCGTGGCGGC 570
Db      167 AlaLeuSerGlnGlyAlaThrProGlnAspLeuAsnThrMetLeuAsnThrValGlyGly 186
```

```
RESULT 12
US-09-952-060-33
; Sequence 33, Application US/09952060
; Patent No. 6733993
; GENERAL INFORMATION:
; APPLICANT: Emini, Emilio A.
```

```

QY 571 CACCAAGCCCGCATGTCAGATGCTGAAGACACCATCAACGAGAGGCCCGCGAGTGGAC 630
   |||||||
Db 187 HieGlnAlaAlaMetGlnMetLeuylsGluThrIleasnGlnGlnAlaAlaGluTrrpAsp 206
QY 631 CGCCTGACACCCCGTGCAGAGCCCGCGCTGGCCCCCGCCAGATGCCCGGACCCCGGGGC 690
   |||||||
Db 207 ArgLeuHisProValHisAlaGlyProIleAlaProGlyGlnMetArgGluProArgGly 226
QY 691 AGCGACATCGCCCGCGCCACCGACACCTTGACGAGACAGATCGCCTGGATGACCAGCAAC 750
   |||||||
Db 227 SerAspIleAlaGlyThrThrSerThrLeuGlnGlnIleGlyTrrpMetThrAsnAsn 246
QY 751 CCCCCCGTGGCCCGTGGCGACATCTACAGCGGTGATCATCTCGGCGCTGAACAAGATC 810
   |||||||
Db 247 ProProIleProValGlyGluIleTyrLysArgTrrpIleIleLeuGlyLeuAsnLysIle 266
QY 811 GTCCGGATGTACAGCCCCGTGAGCATCTTGACATCCGCCAGAGGCCCGCAAGAGCCCTTC 870
   |||||||
Db 267 ValArgMetTyrSerProThrSerIleLeuAspIleArgGlnGlyProLysGluProPhe 286
QY 871 CGCGACTACGTGAGCCGCTTCTCAAGACCTCGCGCCCGAGACAGGCCACCGACGACGTG 930
   |||||||
Db 287 ArgAspTyrValAspArgPheTyrLysThrLeuArgAlaGlnGlnAlaSerGlnGlnVal 306
QY 931 AAGAACTGGATGACCGACACCTCGCTGCTGACAGAACCCCAACCCCGACTGCAAGACCATC 990
   |||||||
Db 307 LysAsnTrrpMetThrGluThrLeuLeuValGlnAsnAlaAsnProAspCysLysThrIle 326
QY 991 CTGCGCGCTCTCGCGCCCGCGCCACCTTGAGAGATGATGACCGCGCTGCAGAGCGGTG 1050
   |||||||
Db 327 LeuLysAlaLeuGlyProAlaAlaThrIleuGlnGlnMetMetThrAlaCysGlnGlyVal 346
QY 1051 GCGCGCCCGCGCCACAAAGGCCCGCGCTGCGCGAGCGGATGAGCCAG--GCCAAAGC 1107
   |||||||
Db 347 GlyGlyProGlyHisLysAlaArgValIleuAlaGlnAlaMetSerGlnValThrAsnSer 366
QY 1108 GTGAACATCATGATGACAGAAAGCAACTTCAAGGGCGCCCGCGCAACGTCAAGTGCCTC 1167
   |||||||
Db 367 AlaThrIleMetMetGlnArgGlyAsnPheArgAsnGlnArgLysThrValLysCysPhe 386
QY 1168 AACTGCGGCAAGAGAGGCCACATCGCCCAAGAACTGCGCGCCCGCCGCAAGAGGCGTGC 1227
   |||||||
Db 387 AsnCysGlyLysValGlyHisIleAlaLysAsnCysArgAlaProArgLysLysGlyCys 406
QY 1228 TGGAACTGCGGCAAGAGGGCCACCATGTAAGACTGCACCGGAGCGCCAGGCCCACTTC 1287
   |||||||
Db 407 TrrpLysCysGlyLysGlnGlyHisGlnMetLysAspCysAsnGlnArgGlnAlaAsnPhe 426
QY 1288 CTGGGCAAGATCTGGCGCCAGCCACAAAGGGCGCCCGCGCAACTTCTGCAGAACCGCAGC 1347
   |||||||
Db 427 LeuGlyLysIleTrrpProSerHisLysGlyArgProGlyAsnPheLeuGlnSerArgPro 446
QY 1348 GAGCCCGCGCCCGCCACCGTGCACCGCGCCCGCCCGCGAGAGCTTCGCGCTC----- 1401
   |||||||
Db 447 Glu-----ProThrAlaProProGluSerPheArgPheGlyGlu 460
QY 1402 GAGAGAACCAACCCCGCGCCCAAGCAGAGGCCCAAGGACCGCGAGCCCTACCGCGAGCCC 1461
   |||||||
Db 461 GluLysThrThrProSerGlnLysGlnGluProIleAspLysGluLeuTyr-----Pro 478
QY 1462 CTGACCGCCCTGCGGAGCCTGTTCCGACGCGCGCCCGCTGAGCCAG 1506
   |||||||
Db 479 LeuAlaSerLeuArgSerLeuPheGlyAsnAspProSerSerGln 493

```

RESULT 13 US-08-127-499A-11

```

; Sequence 11, Application US/08127499A
; Patent No. 5510264
; GENERAL INFORMATION:
; APPLICANT: VAN ALSTYNE, Diane
; APPLICANT: SHARMA, Lawrence Rajendra
; TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED
; TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES

```

```

; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/127,499A
; FILING DATE: 28-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 51916/102/INBI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; US-08-127-499A-11

```

```

Alignment Scores:
Pred. No.: 8.57e-115 Length: 478
Score: 2018.50 Matches: 382
Percent Similarity: 87.86% Conservative: 45
Best Local Similarity: 78.60% Mismatches: 40
Query Match: 71.07% Indels: 19
DB: 1 Gaps: 5

```

US-09-475-704A-4 (1-1509) x US-08-127-499A-11 (1-478)

```

QY 1 ATGGGCGCCCGCGCCAGCATCTGCGGCGAGAACTGGAAGTGGAGAGATCCGC 60
   |||||||
Db 1 MetGlyAlaArgAlaSerValIleuSerGlyGlyLeuAspArgTrrpGluLysIleArg 20
QY 61 CTGCGCCCGCGCGCAAGAAAGCACTACATGCTGAAGCACCTGCTGTGGCGCCGCGAG 120
   |||||||
Db 21 LeuArgProGlyGlyLysLysLysTyrLysLeuLysHisIleValTrrpAlaSerArgGlu 40
QY 121 CTGAGGGGCTTGGCCCTGAACCCCGCGCTGAGACCGCGGAGGCTGCAACAGATC 180
   |||||||
Db 41 LeuGluArgPheAlaValAlaAsnProGlyLeuLeuGluThrSerGluGlyCysArgGlnIle 60
QY 181 ATGAAGCAGCTGACCGCCCGCTGCAGACCGGCAACCGAGAGCTGCGAGCCTGTACAAC 240
   |||||||
Db 61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGlnGluLeuArgSerLeuTyrAsn 80
QY 241 ACCGTGGCAACCTTACTGCTGTCACGCGCGCATGAGGTCGCGACACCAAGAGAGCC 300
   |||||||
Db 81 ThrValAlaThrLeuTyrCysValHisGlnArgIleGluIleLysAspThrLysGluAla 100
QY 301 CTGGAACAAGATCGAGAGAGAGCAAGCAAGTCCACGACAGAACCCAGACCAAGAGAG 360
   |||||||
Db 101 LeuAspLysIleGluGluGlnGlnAsnLysSerLysLysLysAlaGlnGlnAlaAla 120
QY 361 GCCGACGCGC-----AAGGTGAGCCAGAACTACCCCATCGTGCAAGACTGACGAGC 411
   |||||||
Db 121 AspThrGlyHisSerSerGlnValSerGlnAsnTyrProIleValGlnAsnIleGlnGly 140
QY 412 CAGATGTGACACCAAGCCCATGACGCCCGCGACCCCTGAACGCTGGTGGTGAAGTGATCGAG 471
   |||||||
Db 141 GlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTrrpValLysValGlu 160

```

QY 472 GAGAGGCTTCAGCCCCGAGGTGATCCCATGTTACCGCCCTGAGCGAGCGCCACC 531
Db 161 GluLysAlaPheSerProGluValIleProMetPheSerAlaLeuSerGluGluAlaThr 180
QY 532 CCCGAGCACTGAACAGATGTTGAACACCGTGCGCGCCACGAGCGCCCATGACAGTG 591
Db 181 ProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaIleMetGlnMet 200
QY 592 CTGAAGGACACCATCAACGAGAGGCGCGCGAGTGGAACCGCCTGCACCCCGTGACGGC 651
Db 201 LeuLysGluThrIleAsnGluGluAlaIleGluTyrAspArgValHisProValHisAla 220
QY 652 GGGCCCGTGGCCCCCGGCGAGATGCGCGCACCCCGCGGAGCGACATCGCCGCGCCACC 711
Db 221 GlyProIleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThr 240
QY 712 AGCACCCCTGCAGAGAGACATCGCCTGATGACCAACACCCCGCTGCGCCGTGGCGAC 771
Db 241 SerThrLeuGlnGluGlnIleGlyTyrMetThrAsnAsnProProIleProValGlyGlu 260
QY 772 ATCTACAGCGGTGATCATCTCGGCTGGAACAAGATCGTGCGGATGATACGCCCCGTG 831
Db 261 IleTyrLysArgTyrIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProThr 280
QY 832 AGCATCCTTGACATCCGCCAGGCGCCCAAGAGACCCCTTCGCGACTACGTGACCGCTTC 891
Db 281 SerIleLeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhe 300
QY 892 TTCAGAGACCTTGGCGCGGAGAGAGCCACCCAGAGACCGTGAAGAACTGATGACCGAGACC 951
Db 301 TyrLysThrLeuArgAlaGluGlnAlaSerGlnGluValLysAsnTyrMetThrGluThr 320
QY 952 CTGCTGCTGACGAACCGCAACCCCGACTGCAAGACCATCTCGCGCTCTCGGCCCGCGCC 1011
Db 321 LeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyProAla 340
QY 1012 GCCACCCCTGAGAGATGATGACCGCTGCGAGGCGGTGGCGCGCCCGCCACCAAGGCC 1071
Db 341 AlaThrLeuGluGlnMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLysAla 360
QY 1072 CGCGTGTGGCGGAGGATGAGCCAG---GCCAACAGCGTGAACATCATGATGACAGAG 1128
Db 361 ArgValLeuAlaGlnAlaMetSerGlnValThrAsnThrAlaThrIleMetMetGlnArg 380
QY 1129 AGCAACTTCAAGGCGCGCGCGCAAGCTCAAGTGTCTCAACTGCGGCAAGAGGCGCAC 1188
Db 381 GlyAsnPheArgAsnGlnArgLysMetValLysCysPheAsnCysGlyLysGluGlyHis 400
QY 1189 ATGCGCAAGAACTGCGCGCGCGCGCGCAAGAGGCTGCTGGAAGTGGCGCAAGAGGCC 1248
Db 401 ThrAlaArgAsnCysArgAlaProArgLysLysGlyCysTyrLysCysGlyLysGluGly 420
QY 1249 CACCAAGTGAAGACTGCACCGAGCGCGCAAGCTTCTGGGCAAGATC---TGGCCC 1305
Db 421 HisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleCysLeuPro 440
QY 1306 AGCCACAAGGCGCGCGCGCAACTTCTGCAGAACCGGACGAGCGCGCGCCCGCCACC 1365
Db 441 ThrArgGluGlnGlyGlnIlePhePheArgAlaAspGlnSerGln----- 455
QY 1366 GTGCCACCGCGCGCGCGCGAGAGCTTCGCTTCGAGGAGAGACCGCCCGCCCAAG 1425
Db 456 -----GlnProHisHisPhePheArgAlaAspGlnSer----- 466
QY 1426 CAGAGCGCCCAAGAGCCGC 1443
Db 467 GlnGlnProHisGlnLys 472

RESULT 14
US-08-482-847-11
; Sequence 11, Application US/08482847
; Patent No. 5556757

GENERAL INFORMATION:
; APPLICANT: VAN ALSTYNE, Diane
; APPLICANT: SHARMA, Lawrence Rajendra
; TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR
; TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR
; TITLE OF INVENTION: CNS CARRIER, ANTIBODIES THERETO, AND USES THEREOF
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,847
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/127,499
; FILING DATE: 28-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 51916/104/INBI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEEX: 904136
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; US-08-482-847-11

Alignment Scores:
Pred. No.: 8.57e-115 Length: 478
Score: 2018.50 Matches: 382
Percent Similarity: 87.86% Conservative: 45
Best Local Similarity: 78.60% Mismatches: 40
Query Match: 71.07% Indels: 19
DB: 1 Gaps: 5

US-09-475-704A-4 (1-1509) x US-08-482-847-11 (1-478)
QY 1 ATGGGCGCGCGCGCGAGCATCTGCGCGGAGAGCTGGACAAAGTGGAGAAATCCGC 60
Db 1 MetGlyAlaArgAlaSerValLeuSerGlyGlyGluLeuAspArgTyrGluLysIleArg 20
QY 61 CTGGCGCGCGCGCGCAAGAACTACATGCTGAAGCACTGTGTGTGGCCAGCGCGAG 120
Db 21 LeuArgProGlyGlyLysLysLysTyrLysLeuLysHisIleValTyrPalaserArgGlu 40
QY 121 CTGAGGGGCTTGGCCCTGAACCCCGGCTGCTGGAAGACCGCGGAGGGCTGCAAGCAGATC 180
Db 41 LeuGluArgPheAlaValAlaAsnProGlyLeuLeuGluThrSerGluGlyCysArgGlnIle 60
QY 181 ATGAAGCAGCTGACAGCCCGCGCTGCAGACCGGAGACCGAGAGCTGCGAGCGCTGTACAAC 240
Db 61 LeuGlyGlnLeuGlnProSerLeuGlnThrGlySerGluGluLeuArgSerLeuTyrAsn 80
QY 241 ACCGTGGCCACCTGTACTGCGTGCACGCGCGGATCGAGGTCGCGACACCAAGAGGCC 300
Db 81 ThrValAlaThrLeuTyrCysValHisGlnArgIleGluIleLysAspThrLysGluAla 100
QY 301 CTGGAACAAGATGACGAGAGGAGAGCAACAAGTCCAGCAAGAAGACCAGAGGCCAAGAG 360

Db 101 LeuAspLysIleGluGluGlnAlaAsnLysSerLysLysAlaGlnGlnAlaAla 120
 QY 361 GCCGACGGC-----AAGGTAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGC 411
 Db 121 AspThrGlyHisSerSerGlnValSerGlnAsnTyrProIleValGlnAsnIleGlnGly 140
 QY 412 CAGATGCTGCACCAAGCCATCAGCCCCCGCACCCCTGAACGCTGGGTGAAGTGCATGAG 471
 Db 141 GlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaIleValLysValValGlu 160
 QY 472 GAGAAAGCCTTACAGCCCGAGGTGATCCCATGTTCAACCCCTGAGAGAGGGCCGACC 531
 Db 161 GluLysAlaPheSerProGluValIleProMetPheSerAlaLeuSerGluGlyAlaThr 180
 QY 532 CCCACGACCTGAACACGATGTTGAACCGTGGCGGCCACAGCCGCCATGCAGATG 591
 Db 181 ProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMet 200
 QY 592 CTGAAGGACACCATCAACGAGAGGGCCGAGTGGACCGCCTGCACCCCGTGCAGGGC 651
 Db 201 LeuLysGluThrIleAsnGluGlnAlaAlaGluTyrAspArgValHisProValHisAla 220
 QY 652 GGCCCGGTGGCCCCCGCCAGATGCGCGCACCCCGCGCAGCGACATCGCCGCGCCACC 711
 Db 221 GlyProIleAlaProGlyGlnMetArgGluProArgLysSerAspIleAlaGlyThrThr 240
 QY 712 AGCACCTGCAGAGAGATCGCTGATGACCAACCCCGCCGTCGCGCGCGGAC 771
 Db 241 SerThrLeuGlnGluGlnIleGlyTyrMetThrAsnAsnProProIleProValGlyGlu 260
 QY 772 ATCTACAGCGGTGATCATCTGGGCTGAACAAGATCGTGGATGTACAGCCCGGTG 831
 Db 261 IleTyrLysArgTyrIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProThr 280
 QY 832 AGCATCTGCAGATCCGCCAGGGGCCCAAGAGAGCCCTTCGCGACATCGTGCAGCCGCTTC 891
 Db 281 SerIleLeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhe 300
 QY 892 TTCAAGACCTTGCGCGCGCAGAGGCCACCCAGACGCTGAAGAACTGCATGCAGAGACC 951
 Db 301 TyrLysThrLeuArgAlaGluGlnAlaSerGlnGluValLysAsnTyrMetThrGluThr 320
 QY 952 CTGCTGGTGCAGAAACGCCAACCCCGCATGCAAGACCATCTCGCGCTCTCGGCCCGCCG 1011
 Db 321 LeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyProAla 340
 QY 1012 GCCACCTGAGAGATGATGACCGCTGCCAGGCGCTGGGGCGGCCGCCACAGAGCC 1071
 Db 341 AlaThrLeuGluGluMetMetThrAlaCysGlnGlyValGlyProGlyHisLysAla 360
 QY 1072 CGCGTGTGGCCGAGGCGATGAGCCAG--GCCAACGCGTGAACATCATGATGCAGAG 1128
 Db 361 ArgValLeuAlaGlnAlaMetSerGlnValThrAsnThrAlaThrIleMetMetGlnArg 380
 QY 1129 AGCAACTTCAAGGGCGCCCGCGCAACGTCAAGTCTTCAACTGCGGCAAGAGGGCCAC 1188
 Db 381 GlyAsnPheArgAsnGlnArgLysMetValLysCysPheAsnCysGlyLysGluGlyHis 400
 QY 1189 ATCGCCAGAACTGCGCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGGCGCAAGAGGGC 1248
 Db 401 ThrAlaArgAsnCysArgAlaProArgLysLysGlyCysTyrLysCysGlyLysGluGly 420
 QY 1249 CACCAAGTGAAGACTGCACCGAGCGCCAGCCACTTCTGGGCAAGATC--TGGCCC 1305
 Db 421 HisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleCysLysLeuPro 440
 QY 1306 AGCCACAAAGGGCGCGCCGCGCAACTTCTGCAGAACCGCGAGAGCCCGCGCCGCCACC 1365
 Db 441 ThrArgGluGlyGlnGlyIlePhePheArgAlaAspGlnSerGln----- 455
 QY 1366 GTGCCACCGCGCCCGCGAGAGCTTCCGCTTCGAGAGAGACCCCGCGCCCGCCAG 1425

Db 456 -----GlnProHisIlePhePheArgAlaAspGlnSer----- 466
 QY 1426 CAGAGCCCAAGAGACCGC 1443
 Db 467 GlnGlnProHisGlnLys 472
 RESULT 15
 US-08-392-794A-2
 ; Sequence 2, Application US/08392794A
 ; Patent No. 6025141
 ; GENERAL INFORMATION:
 ; APPLICANT: HU, Yu-Wen
 ; TITLE OF INVENTION: IMMUNOFLOURESCENCE ASSAY FOR THE DETECTION OF
 ; TITLE OF INVENTION: ANTIBODIES USING RECOMBINANT ANTIGENS IN INSOLUBLE FORM
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BROWDY AND NEIMARK
 ; STREET: 419 Seventh Street, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20004
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/392,794A
 ; FILING DATE: 09-DEC-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/164,789
 ; FILING DATE: 10-DEC-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: NEIMARK, SHERIDAN
 ; REGISTRATION NUMBER: 20,520
 ; REFERENCE/DOCKET NUMBER: HU=4A
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-628-5197
 ; TELEFAX: 202-737-3528
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 437 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-392-794A-2
 Alignment Scores:
 Pred. No.: 2,41e-114 Length: 437
 Score: 2011.00 Matches: 374
 Percent Similarity: 93.14% Conservative: 33
 Best Local Similarity: 85.58% Mismatches: 26
 Query Match: 70.81% Indels: 4
 DB: 3 Gaps: 2
 US-09-475-704A-4 (1-1509) x US-08-392-794A-2 (1-437)
 QY 1 ATGGGCGCGCGCCAGCATCTGCGCGCGAGAGCTGGAAGTGGAGAGATCCGC 60
 Db 1 MetGlyAlaArgAlaSerValLeuSerGlyGlyGluLeuAspArgTyrGluLysIleArg 20
 QY 61 CTGCGCGCGCGCGCAAGAGCACTACATGCTGAAGCACCTGCTGGGCGCGAG 120
 Db 21 LeuArgProGlyGlyLysLysTyrLysLysLysHisIleValTyrPalaSerArgGlu 40
 QY 121 CTGAGGCGCTTGGCTTGAACCCCGCGCTGTCGAGACCGCGAGGCTGCAGACAGATC 180
 Db 41 LeuGluArgPheAlaValAlaAsnProGlyLeuLeuGluThrSerGluGlyCysArgGlnIle 60
 QY 181 ATGAAGCAGCTGCAGCCCGCGCTGCAGACCGGCAACGAGAGCTGCGACGCTGTACAAC 240

Db 61 LeuGlyInLeuGlnProSerLeuGlnThrGlySerGlnGluLeuArgSerLeuTyrAsn 80
QY 241 ACCGTGGCCACCTGTACTGCGTGACGCCGCGCATCGAGGTCCGCGACACCAAGAGGCC 300
Db 81 ThrValAlaThrLeuTyrCysValHisGlnArgIleGluIleLysAspThrLysGluAla 100
QY 301 CTGGAACAAGATCGAGGAGAGCAGAACAGTCCACGACAAGACCCAGCAGGCCAAGAG 360
Db 101 LeuAspLysIleGlnGluGlnGlnAsnLysSerLysLysAlaGlnGlnAlaAla 120
QY 361 GCCGAGGC-----AAGGTAGCCAGAACTACCCCATCGTGCAAGACCTGAGGCC 411
Db 121 AspThrGlyHisSerAsnGlnValSerGlnAsnTyrProIleValGlnAsnIleGlnGly 140
QY 412 CAGATGGTGCAACGACGATCAGCCCCCGCACCTGAACGCTGGGTGAAGTGATCGAG 471
Db 141 GlnMetValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValValGlu 160
QY 472 GAGAAGGCTTCAGCCCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGCCACC 531
Db 161 GluLysAlaPheSerProGluValIlePrometPheSerAlaLeuSerGlnGlyAlaThr 180
QY 532 CCCCAAGACCTGAACAGATGTTGAACACCGTGGCGGCCACCAAGGCCCATGCAATG 591
Db 181 ProGlnAspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaIleMetGlnMet 200
QY 592 CTGAAGACACCATCAACGAGGAGGCCCGCGAGTGGACCGCTGCACCCCGTGCAAGCC 651
Db 201 LeuLysGluThrIleAsnGlnGluAlaIleGluTrpAspArgValHisProValHisAla 220
QY 652 GGGCCCGTGGCCCCCGGCAGATGCGCGGACCCCGCGGCAAGCAGCATCGCCGGCGCCACC 711
Db 221 GlyProIleAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThr 240
QY 712 AGCACCTTGCAAGAGCAGATCGCTGATGACAGCAACCCCGCGTGGCGGCGAC 771
Db 241 SerThrLeuGlnGlnGlnIleGlyTrpMetThrAsnAsnProIleProValGlyGlu 260
QY 772 ATCTACAAGCGGTGATCATCTGGGCTTGAACAAGATCGTGGATGTACAGCCCCGTG 831
Db 261 IleTyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProThr 280
QY 832 AGCATCTTGACATCGGCCAGAGGGCCCCCAAGAGAGCCCTTCGCGACTACGTGACCGCTTC 891
Db 281 SerIleLeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhe 300
QY 892 TTCAAGACCTTGCGCGCGCGAGCAGAGGCCACCCAGAGCGTGAAGAACTGATGACCGAGACC 951
Db 301 TyrLysThrLeuArgAlaGlnGlnAlaSerGlnGlnValLysAsnTrpMetThrGluThr 320
QY 952 CTGCTGGTGCAAGACGCCCAACCCCGCATGCAAGACCATCTGCGCGCTCTGGCCCGGC 1011
Db 321 LeuLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyProAla 340
QY 1012 GCCACCTTGAGGAGATGATGACCGCTGCCAGGGCGGTGGCGGCCCGGCACAGAGGCC 1071
Db 341 AlaThrLeuGlnGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLysAla 360
QY 1072 CGCGTCTGGCCGAGCGCATGAGCCAG--GCCAACAGCGTGAACATCATGATGCAAG 1128
Db 361 ArgValLeuAlaGlnAlaMetSerGlnValThrAsnSerAlaThrIleMetMetGlnArg 380
QY 1129 AGCACTTCAAGGGCCCCCGGCGCAAGTCAAGTCTTCAACTGCGGCAAGAGAGGCCAC 1188
Db 381 GlnAsnPheArgAsnGlnArgLysIleValLysCysPheAsnCysGlyLysGlnGlyHis 400
QY 1189 ATGCCAAGAAGCTGCGCGCCCCCGGCAAGAGGGGTGCTGGAAGTCCGCAAGAGAGGCC 1248
Db 401 ThrAlaArgAsnCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGlnGly 420
QY 1249 CACCAAGTGAAGACTGACCGGAGCGCCAGGCCAACTTCTGGGCAAGATC 1299

Db 421 HisGlnMetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIle 437

Search completed: September 1, 2005, 23:19:54
Job time : 48.261 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 1, 2005, 22:09:18 ; Search time 386.088 Seconds
(without alignments)
3078.466 Million cell updates/sec

Title: US-09-475-704A-4
Perfect score: 2840
Sequence: 1 atggcgccgcgcgcacgat.....gcggcccccctgagccagtaa 1509

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1774312 seqs, 393823214 residues
Total number of hits satisfying chosen parameters: 3548624

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09475704/runat_01092005_160752_19548/app_query.fasta_1.3342
-DB=Published_Applications_AA -QFMT=fastran -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09475704_@CGN_1_1_233_@runat_01092005_160752_19548
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOB=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2661	93.7	502	10 US-09-899-575-22	Sequence 22, Appl
2	2467	86.9	492	18 US-10-780-507-97	Sequence 97, Appl
3	2464	86.8	492	16 US-10-332-413-2	Sequence 2, Appl
4	2464	86.8	492	18 US-10-494-131-36	Sequence 36, Appl
5	2457	86.5	492	9 US-09-991-258-5	Sequence 5, Appl
6	2457	86.5	492	18 US-10-929-234-5	Sequence 5, Appl
7	2457	86.5	492	16 US-10-332-413-8	Sequence 8, Appl
8	2456.5	86.5	493	18 US-10-780-507-98	Sequence 98, Appl
9	2450	86.3	492	18 US-10-780-507-96	Sequence 96, Appl
10	2440	85.9	492	14 US-10-339-217-109	Sequence 109, App
11	2430	85.6	492	14 US-10-339-217-110	Sequence 110, App
12	2421.5	85.3	491	14 US-10-339-217-140	Sequence 140, App
13	2413.5	85.0	491	14 US-10-339-217-143	Sequence 143, App
14	2387	84.0	508	14 US-10-339-217-105	Sequence 105, App
15	2385.5	84.0	491	14 US-10-339-217-104	Sequence 104, App
16	2375.5	83.6	497	14 US-10-339-217-107	Sequence 107, App
17	2353.5	82.9	487	14 US-10-339-217-106	Sequence 106, App
18	2348	82.7	496	14 US-10-339-217-108	Sequence 108, App
19	2343	82.5	492	10 US-09-899-575-17	Sequence 17, Appl
20	2263	79.7	500	18 US-10-780-507-47	Sequence 47, Appl
21	2258.5	79.5	1228	18 US-10-494-131-30	Sequence 30, Appl
22	2258	79.5	500	14 US-10-339-217-18	Sequence 18, Appl
23	2251	79.3	442	18 US-10-494-131-32	Sequence 32, Appl
24	2247.5	79.1	494	14 US-10-339-217-126	Sequence 126, App
25	2246	79.1	500	14 US-10-223-172A-42	Sequence 42, Appl
26	2246	79.1	500	14 US-10-224-999A-3483	Sequence 3483, Ap
27	2246	79.1	500	18 US-10-780-507-48	Sequence 48, Appl
28	2240.5	78.9	497	16 US-10-800-023-21	Sequence 21, Appl
29	2238	78.8	1350	10 US-09-952-060-35	Sequence 35, Appl
30	2238	78.8	1350	16 US-10-380-641-35	Sequence 35, Appl
31	2238	78.8	1350	17 US-10-636-730-35	Sequence 35, Appl
32	2237	78.8	512	17 US-10-634-165-8	Sequence 8, Appl
33	2234	78.7	500	18 US-10-780-507-49	Sequence 49, Appl
34	2232	78.6	557	14 US-10-339-217-144	Sequence 144, App
35	2226	78.4	500	16 US-10-325-468-8	Sequence 8, Appl
36	2225.5	78.4	494	14 US-10-339-217-100	Sequence 100, App
37	2223	78.3	500	14 US-10-097-534-32	Sequence 32, Appl
38	2223	78.3	500	14 US-10-059-271-85	Sequence 85, Appl
39	2223	78.3	500	14 US-10-102-622-4	Sequence 4, Appl
40	2223	78.3	500	16 US-10-325-468-34	Sequence 34, Appl
41	2223	78.3	500	17 US-10-844-658-6	Sequence 6, Appl
42	2221	78.2	500	9 US-09-968-355-26	Sequence 26, Appl
43	2221	78.2	583	9 US-09-968-355-17	Sequence 17, Appl
44	2220.5	78.2	499	17 US-10-844-658-2	Sequence 2, Appl
45	2220	78.2	498	16 US-10-325-468-22	Sequence 22, Appl

ALIGNMENTS

RESULT 1
US-09-899-575-22
; Sequence 22, Application US/09899575
; Publication NO. US20030223961A1
; GENERAL INFORMATION:
; APPLICANT: Zur Megele, Jan
; APPLICANT: Barnett, Susan W.
; APPLICANT: Egnelbrecht, Susan
; APPLICANT: van Rensburg, Estrelita Janse
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP01631.102
; CURRENT APPLICATION NUMBER: US/09/899,575
; CURRENT FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 09/475,704
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 502

```
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-899-575-22

Alignment Scores:
Pred. No.:      9.17e-129      Length:      502
Score:          2661.00      Matches:      502
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches:  0
Query Match:    93.70%      Indels:      0
DB:             10      Gaps:      0

US-09-475-704A-4 (1-1509) x US-09-899-575-22 (1-502)

QY      1 ATGGGCGCGCGCGCGCATCTGCGCGCGGAGAAGCTGGACAAGTGGGAGAAAGATCCGC 60
Db      1 MetGlyAlaArgAlaSerIleLeuArgGlyGluLysLeuAspLysIleArg 20
QY      61 CTGCGCGCGCGCGCGCAAGAAAGCACTACATGCTGAAGCACTGGTGTGGGCCAGCGCGGAG 120
Db      21 LeuArgProGlyGlyLysLysHisTyrMetLeuLysHisLeuValIleArg 40
QY      121 CTGAGAGGCTTGGCGCTGAAACCCCGGCTGTGAGACCGCGAGGGCTGCAAGAGATC 180
Db      41 LeuGluGlyPheAlaLeuAsnProGlyLeuLeuGluThrAlaGluGlyCysLysGlnIle 60
QY      181 ATGAAGCAGCTGCAGCGCGCGCGCTGCAGACCGGCACGAGGAGCTGCGCAGCTGTACAAC 240
Db      61 MetLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluGluLeuArgSerLeuTyrAsn 80
QY      241 ACCGTGGCCACCTCTGTACTGCGTGACGCGCGGATCGAGGTCGCGCACCAAGAGGCGC 300
Db      81 ThrValAlaThrLeuTyrCysValHisAlaGlyIleGluValArgAspThrLysGlnAla 100
QY      301 CTGGAACAAGATCGAGGAGGAGCAGAAACAAGTCCACAGAGAAGACCAAGGCCAAGAGAG 360
Db      101 LeuAspLysIleGluGluGluGlnAsnLysSerGlnGlnLysThrGlnGlnAlaLysGln 120
QY      361 GCCGAGCGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAAGGGCCAGATGGTG 420
Db      121 AlaAspGlyLysValSerGlnAsnTyrProIleValGlnAsnLeuGlnGlyGlnMetVal 140
QY      421 CACCAAGGCCATCAGCGCGCGCGCACCCCTGAACGCTGGTGAAGGTGATCGAGGAGAGGCC 480
Db      141 HisGlnAlaIleSerProArgThrLeuAsnAlaIleThrValLysValIleGluGluLysAla 160
QY      481 TTCAGCGCGCGAGGTGATCCCATGTTTCAACGCGCTGAGCGAGGGCGGCCACCGCCAGGAC 540
Db      161 PheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaIleThrProGlnAsp 180
QY      541 CTGAACACGATGTTGAACAACCGTGGGCGGCGCACAGGCGCGCATGACATGCTGAAGAAC 600
Db      181 LeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLysAsp 200
QY      601 ACCATCAACGAGAGGCGCGCGAGTGGGACCGCTGCAACCGCGTGCAGCGCGCGCGCGTG 660
Db      201 ThrIleAsnGluGluAlaIleGluIleTrpAspArgLeuHisProValGlnAlaGlyProVal 220
QY      661 GCCCGCGCGCAGATGCGCGGACCGCGCGGCGAGCAGACATGCGCGCGCGCCAGCACCGCTG 720
Db      221 AlaProGlyGlnMetArgAspProArgLysSerAspIleAlaGlyAlaIleThrSerThrLeu 240
QY      721 CAGAGCAGATGCGCTGGATGACCAAGCAACCGCGCTGCGCGCGTGGGCGACATCTACAAG 780
Db      241 GlnGluGlnIleAlaIleArgMetThrSerAsnProProValProValGlyAspIleTyrLys 260
QY      781 CGGTGATCATCTGGGCGCTGAACAAGATCGTGCGATGTACAGCGCGCTGAGCATCTTG 840
Db      261 ArgTyrIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProValSerIleLeu 280
QY      841 GACATCGCCAGGCGCGCAAGAGCGCTTCCGCGACTACGTGACCGCTTCTTCAAGACC 900
Db      281 AspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhePheLysThr 300
```

```
QY      901 CTGCGCGCGGAGGCGCACCGCAGAGCTGAAGAAGTGTGATGACCGAGACCTGTGTG 960
Db      301 LeuArgAlaGluGlnAlaIleThrGlnAspValLysAsnIleTrpMetThrGluThrLeuVal 320
QY      961 CAGAAAGCCAAACCGCGACTGCAAGACCATCTGCGCGCTGTGCGCGCGCGCGCACCGCTG 1020
Db      321 GlnAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeuGlyProGlyAlaIleThrLeu 340
QY      1021 GAGAGATGATGACCGCTGCGCAGGCGCGTGGGCGCGCGCGCGCGCGCAAGGCGCGCTGTG 1080
Db      341 GluGluMetMetThrAlaCysGlnGlnGlyValGlyGlyProGlyHisLysAlaArgValLeu 360
QY      1081 GCCGAGCGGATGAGCCAGGCGCAACAGCGTGAACATCATGATGACGAAGAGCACTTCAAG 1140
Db      361 AlaGluAlaMetSerGlnAlaAsnSerValAsnIleMetMetGlnLysSerAsnPheLys 380
QY      1141 GCGCGCGCGCGCAACGTCAAGTGTCTCAACTGCGCGCAAGAGGCGCGCACATCGCGCAAGAAC 1200
Db      381 GlyProArgArgAsnValLysCysPheAsnCysGlyLysGluGluGlnHisIleAlaLysAsn 400
QY      1201 TGC CGCGCGCGCGCGCAAGAGGCGTGTGGAAGTGCAGGAGGCGCGCACAGATGAAG 1260
Db      401 CysArgAlaProArgLysLysGlyCysTyrLysCysGlyLysGluGluGlnHisIleMetLys 420
QY      1261 GACTGCACCGAGCGCGCAGGCGCAACTTCTGCGCGCAAGATCTGCGCGCACAGGCGCGCGC 1320
Db      421 AspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSerHisLysGlyArg 440
QY      1321 CCGCGCAACTTCTGCGAAGACCGCAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 1380
Db      441 ProGlyAsnPheLeuGlnAsnArgSerGluProAlaAlaProThrValProThrAlaPro 460
QY      1381 CCGCGCGAGAGCTTCCGCTTGCAGAGAGACCAACCGCGCGCGCGCGCGCGCGCGCGCGCGC 1440
Db      461 ProAlaGluSerPheArgPheGluGluThrThrProAlaProLysGlnGluProLysAsp 480
QY      1441 CGCGAGCCCTACCGCGGAGCGCGCTGACCGCGCTGCGGAGCGCTGTGCGAGCGCGCGCGCTG 1500
Db      481 ArgGluProTyrArgGluProLeuThrAlaLeuArgSerLeuPheGlySerGlyProLeu 500
QY      1501 AGCCAG 1506
Db      501 SerGln 502

RESULT 2
US-10-780-507-97
; Sequence 97, Application US/10780507
; Publication No. US20050137387A1
; GENERAL INFORMATION:
; APPLICANT: MULLINS, James I.
; APPLICANT: RODRIGO, Allen G.
; APPLICANT: LEARN, Gerald H.
; APPLICANT: LI, Fusheng
; APPLICANT: NICKLE, David C.
; APPLICANT: JENSEN, Mark A.
; TITLE OF INVENTION: ANCESTRAL AND COT VIRAL SEQUENCES, PROTEINS AND IMMUNOGENIC COMPO
; FILE REFERENCE: 16336-001320US
; CURRENT APPLICATION NUMBER: US/10/780,507
; PRIOR FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US 10/204,204
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US 60/447,586
; PRIOR FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 97
; LENGTH: 492
; TYPE: PRT
```

```

; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Least squares center of tree reconstruction of clade C gag protei
; US-10-780-507-97

```

Alignment Scores:

Pred. No.:	8.41e-119	Length:	492
Score:	2467.00	Matches:	468
Percent Similarity:	95.62%	Conservative:	12
Best Local Similarity:	93.23%	Mismatches:	10
Query Match:	86.87%	Indels:	2
DB:	18	Gaps:	2

US-09-475-704A-4 (1-1509) x US-10-780-507-97 (1-492)

```

QY 1 ATGGGGCGCCGCGCCAGCATCTGCGCGCGGAGAGCTGACAAAGTGGAGAAAGATCCGC 60
Db 1 MetGlyAlaArgAlaSerIleLeuArgGlyGlyLysLeuAspThrTrpGluLysIleArg 20
QY 61 CTGCGCGCGCGCGGCAAGAACACTACATGCTGAAGCACCTGTGTGGCCAGCCGGAG 120
Db 21 LeuArgProGlyGlyLysLysHisTyrMetLeuLysHisLeuValTrpAlaSerArgGlu 40
QY 121 CTGAGGGCTTCGCGCTGAACCCCGCGCTGCTGAGACCGCGAGGCGCTGCAAGCAGATC 180
Db 41 LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrSerGluGlyCysLysGlnIle 60
QY 181 ATGAAGCAGCTGCAGCCCGCTGTCAGACCGGCACCGAGAGCTGCAGCCTGTACAC 240
Db 61 MetLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluGluLeuArgSerLeuTyrAsn 80
QY 241 ACCGTGGCCACCCCTGTAAGTGCAGCCCGCATCGAGTCCGCGACACCAAGAGGCC 300
Db 81 ThrValAlaThrLeuTyrCysValHisGluLysIleGluValArgAspThrLysGlnAla 100
QY 301 CTGACAAGATCGAGAGAGAGCAGAAAGTCCCGACAGAGAACCCAGAGGCCAAGAG 360
Db 101 LeuAspLysIleGluGluGlnGlnAsnLysSerGlnGlnLysThrGlnGlnAlaGlnAla 120
QY 361 GCCGACGCGAAGGTGAGCCAGAACTACCCCATGCTGTCAGAACCTGACGGCCAGATGGTG 420
Db 121 AlaAspGlyLysValSerGlnAsnTyrProIleValGlnAsnLeuGlnGlyGlnMetVal 140
QY 421 CACGAGGCCATCAGCCCGCGCACCCCTGAACGCGCTGGTGAAGTGATCGAGAGAGGCC 480
Db 141 HisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValIleGluGluLysAla 160
QY 481 TTCAGCCCCGAGGTGATCCCATGTTCAACCGCCCTGACGCGAGGGCGGCCACCCCGCAGAC 540
Db 161 PheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGlnAsp 180
QY 541 CTGAACACGATGTTGAACACCGTGGGGCGGCACAGGCGGCCATGACATGCTGAAGAC 600
Db 181 LeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLysAsp 200
QY 601 ACCATCAACGAGAGAGCGCGCGAGTGGGACCGCTGCACCCCGCTGACAGCGCGCGCGCTG 660
Db 201 ThrIleAsnGluGlnAlaAlaGluTrpAspArgLeuHisProValHisAlaGlyProVal 220
QY 661 GCGCGCGCGCGAGATGCGCGACCCCGCGCGCGAGCAGCATCGCGCGCGCGCACCGACCTG 720
Db 221 AlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerThrLeu 240
QY 721 CAGAGCAGATCGCGCTGATGACCAAGAACCCCGCGCTGCCGTGGGCGACATCTACAG 780
Db 241 GlnGluGlnIleAlaIleTrpMetThrSerAsnProProValProValGlyAspIleTyrLys 260
QY 781 CGGTGATCATCTGGGCTTGAACAAGATCGTGGGATGTACAGCCCGCTGAGCATCTG 840
Db 261 ArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProValSerIleLeu 280
QY 841 GACATCCGCCAGGGCGCCCAAGAGGCCCTTCCGCGACTACGTGAGCCGCTTCTTCAAGACC 900

```

```

Db 281 AspIleLysGlnGlyProLysGluProPheArgAspTyrValAspArgPheLysThr 300
QY 901 CTGCGCGCGCGAGAGGCCACCGACGTAAGAACTGGATGACCGAGACCTGCTGGTG 960
Db 301 LeuArgAlaGluGlnAlaThrGlnAspValLysAsnTrpMetThrAspThrLeuVal 320
QY 961 CAGAACCCCAACCCCGACTGCAAGACCATCTGCGCGCTCTCGGCCCGCGCGCACCCCTG 1020
Db 321 GlnAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeuGlyProGlyAlaThrLeu 340
QY 1021 GAGGAGATGATGACCGCGCTGCGAGGCGCTGGGCGCGCGCGCGCACAGGCGCGCTG 1080
Db 341 GlnGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLysAlaArgValLeu 360
QY 1081 GCGGAGCGGATGAGCCAGGCCAAGCGTGAACATCATGATGACAGAAAGCAACTTCAAG 1140
Db 361 AlaGluAlaMetSerGlnAlaAsnAsnThrAsnIleMetMetGlnArgSerAsnLys 380
QY 1141 GCGCGCGCGCGCAACGTCAGTGTCTTCAACTGCGCGCAAGAGGCGCACATCGCCCAAGAC 1200
Db 381 GlyProLysArgIleValLysCysPheAsnCysGlyLysGlnGlyHisIleAlaArgAsn 400
QY 1201 TGCGCGCGCGCGCGCAAGAGGCGCTGCTGGAAGTCCGCGCAAGAGGCGCACAGATGAAG 1260
Db 401 CysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGlnGlyHisGlnMetLys 420
QY 1261 GACTGCACCGAGCGCGCAAGGCCAATTCTTGGGCAAGATCTGGGCCAGCCACAAGGCGCGC 1320
Db 421 AspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSerHisLysGlyArg 440
QY 1321 CCGGCGCACTTCTTGCAGAACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Db 441 ProGlyAsnPheLeuGlnSerArgProGlu-----ProThrAlaPro 454
QY 1381 CCGGCGGAGAGCTTCCGCTTGCAGAGAGACCAACCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
Db 455 ProAlaGluSerPheArgPheGluGluThrThrProAlaProLysGlnGluProLysAsp 474
QY 1441 CCGGAGCCCTACCGCGAGCGCGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 1500
Db 475 -----ArgGluProLeuThrSerLeuLysSerLeuPheGlySerAspProLeu 490
QY 1501 AGCCAG 1506
Db 491 SerGln 492

```

RESULT 3

US-10-332-413-2

; Sequence 2, Application US/10332413

; Publication No. US20040116660A1

; GENERAL INFORMATION:

; APPLICANT: Johnstone, Robert Edward

; APPLICANT: Swannstrom, Ronald Ivar

; APPLICANT: Morris, Lynn

; APPLICANT: Karim, Salim Abdool

; APPLICANT: Williamson, Carolyn

; TITLE OF INVENTION: Process for the Selection of HIV-1 Subtype C Isolates, Selected H

; TITLE OF INVENTION: Subtype Isolates, Their Genes and Modifications and Derivatives (T

; FILE REFERENCE: 45669-281993

; CURRENT APPLICATION NUMBER: US/10/332,413

; CURRENT FILING DATE: 2003-09-12

; PRIOR APPLICATION NUMBER: PCT/IB01/01208

; PRIOR FILING DATE: 2001-07-09

; PRIOR APPLICATION NUMBER: US 60/216,995

; PRIOR FILING DATE: 2000-07-07

; PRIOR APPLICATION NUMBER: ZA 2000/3437

; PRIOR FILING DATE: 2000-07-10

; PRIOR APPLICATION NUMBER: ZA 2000/4924

; PRIOR FILING DATE: 2000-09-15

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

: LENGTH: 492
: TYPE: PRT
: ORGANISM: Human immunodeficiency virus type 1
US-10-332-413-2

Alignment Scores:
Pred. No.: 1.2e-118 Length: 492
Score: 2464.00 Matches: 466
Percent Similarity: 95.62% Conservative: 14
Best Local Similarity: 92.83% Mismatches: 12
Query Match: 86.76% Indels: 10
DB: 16 Gaps: 2

US-09-475-704A-4 (1-1509) x US-10-332-413-2 (1-492)

QY 1 ATGGGCGCCGCGCCAGCATCTGCGCGCGGAGAGCTGGACAAGTGGGAGAAGATCCGC 60
Db 1 MetGlyAlaArgAlaSerIleLeuArgGlyGluLeuAspLysIleArg 20
QY 61 CTGCGCCCGCGCGGCAAGACACTACATGCTGAAGCACCCTGTGTGGCCAGCCGCGAG 120
Db 21 LeuArgProGlyGlyLysLysHisIleTyrMetLeuLysHisIleValTrpAlaSerArgGlu 40
QY 121 CTGAGGGCTTCGCGCTGAACCCCGGCTGTGAGACCGCGGAGGCTGCAAGCATC 180
Db 41 LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrSerGluGlyCysLysGlnIle 60
QY 181 ATGAGCAGCTGCAGCCCGCCCTGCAGACCGGACAGAGAGCTGCGCAGCCTGTACAAC 240
Db 61 MetLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluGluLeuLysSerLeuTyrAsn 80
QY 241 ACCGTGGCCACCTCTACTGCTGTCAGCGCGGACATGAGTCCGCGACCAAGAGAGCC 300
Db 81 ThrValAlaThrLeuTyrCysValHisGluLysIleGluValArgAspThrLysGluAla 100
QY 301 CTGACAAGATCGAGAGGAGAGACAAGTCCCGACAGAGACCCAGCGGCAAGAGAG 360
Db 101 LeuAspLysIleGluGluGlnGluAsnLysCysGlnGlnLysThrGlnGlnAlaLysAla 120
QY 361 GCGGACGGCAAGGTAGCCAGAACCTACCCCATCGTGCAGAACCTGCAGGCGCAGATG 420
Db 121 AlaAspGlyLysValSerGlnAsnTyrProIleValGlnAsnLeuGlnGlyGlnMetVal 140
QY 421 CACCAAGCCATCAGCCCGCGCACCCCTGAACGCGTGGTGAAGTGATCGAGAGAGGCC 480
Db 141 HisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValIleGluGluLysAla 160
QY 481 TTCAGCCCGGAGGTGATCCCATATGTTCAACGCGCTGAGCGAGGCGGCCACCCCGAGAC 540
Db 161 PheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGlnAsp 180
QY 541 CTGAACAAGATGTTGAACACCGTGGCGGCGCACAGCGCCGATGAGATGCTGAAGAC 600
Db 181 LeuAsnThrMetLeuAsnThrValGlyGlnHisGlnAlaAlaMetGlnMetLeuLysAsp 200
QY 601 ACCATCAACGAGAGGCGCGCGAGTGGGAGCCGCTGCACCCCGTGCAGGCGCGCCCGTG 660
Db 201 ThrIleAsnGluGluAlaAlaGluTrpAspArgValHisProValHisAlaGlyProIle 220
QY 661 GCCCGCGCGCAGATGCGCGCACCCCGCGGCGACGACATCGCCGCGGCCACAGCACCTG 720
Db 221 AlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerThrLeu 240
QY 721 CAGAGCAGATCGCTTGATGACCAAGCAACCCCGCTGCGCCCGTGGCGGACATCTACAAG 780
Db 241 GlnGluGlnIleAlaIleTrpMetThrSerAsnProProIleProValGlyAspIleTyrLys 260
QY 781 CGGTGATCATCTGGGCTTGAACAAGATCGTGGGATGTACAGCCCGTGAAGCATCTG 840
Db 261 ArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProValSerIleLeu 280
QY 841 GACATCCGCGCAGGCGCCCAAGAGAGCCCTTCGCGACTACGTGAGCCGCTTCTCAAGACC 900

Db 281 AspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhePheLysThr 300
QY 901 CTGGCGCGCGAGAGGCCACCCAGACGCTGAAGAACTGATGACCGAGACCTGTGCTGTG 960
Db 301 LeuArgAlaGluGlnAlaThrGlnGluValLysAsnTyrMetThrAspThrLeuLeuVal 320
QY 961 CAGAACGCCAACCCGACTGCAAGACCATCTCGCGCGCTCTCGGCCCGCGCCACCCCTG 1020
Db 321 GlnAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeuGlyProGlyAlaThrLeu 340
QY 1021 GAGAGATGATGACCGCTGCGCAGGCGGTGGCGGCCCGGCCACAAGGCCCGCTGCTG 1080
Db 341 GluGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisIleLysAlaArgValLeu 360
QY 1081 GCCGAGCGGATGAGCCAGCGCCACAGCGTGAACATCATGATGCAGAGCAACTTCAAG 1140
Db 361 AlaGluAlaMetSerGlnAlaAsnSerGlyAsnIleMetMetGlnArgSerAsnPheLys 380
QY 1141 GCGCCCGCGCGCAACGTCAAGTCTTCAACTGCGGCAAGAGGCGCCACATCGCCAAGAAC 1200
Db 381 GlyProArgArgIleValLysCysPheAsnCySGlyLysGluGlyHisIleAlaArgAsn 400
QY 1201 TGCCGCGCGCGCGCCCAAGAGGGCTGCTGGAAGTCCGGCAAGAGGCGCCACCATGAAG 1260
Db 401 CysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGluGlyHisIleGlnMetLys 420
QY 1261 GACTGCACCGAGCGCCAGCGCCAACTTCTGGGCAAGATCTGGCGCCAGCCACAAGGCGCG 1320
Db 421 AspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSerHisLysGlyArg 440
QY 1321 CCGGCGCACTTCTGCAAGAACCGCAGCGAGCGCGCGCGCCCGCCACCGTCCCGCGCCCC 1380
Db 441 ProGlyAsnPheLeuGlnAsnArgProGlu-----ProThrAlaPro 454
QY 1381 CCGGCGGAGAGCTTCGCTTCGAGGAGAGACCCCGCGCCCAAGCAGAGAGCCCAAGGAC 1440
Db 455 ProAlaGluSerPheArgPheGluGluThrThrProAlaProLysGlnGluProIleGlu 474
QY 1441 CGGAGGCCCTACCGCGAGCGCCCTGACCGCGCTGCGGACGCTGTTCGGCAGCGGCCCCCTG 1500
Db 475 -----ArgGluProLeuThrSerLeuLysSerLeuPheGlySerAspProLeu 490
QY 1501 AGCCAG 1506
Db 491 SerGln 492

RESULT 4

US-10-494-131-36
: Sequence 36, Application US/10494131
: Publication No. US20050176929A1
: GENERAL INFORMATION:
: APPLICANT: Williamson, Carolyn
: APPLICANT: Van Harmelen, Joanne Heidi
: APPLICANT: Gray, Clive Maurice
: APPLICANT: Bourn, William
: APPLICANT: Karim, Salim Abdool
: TITLE OF INVENTION: HIV-1 Subtype Isolate Regulatory/Accessory Genes, and
: TITLE OF INVENTION: Modifications and Derivatives Thereof
: FILE REFERENCE: 45669-300571
: CURRENT APPLICATION NUMBER: US/10/494,131
: CURRENT FILING DATE: 2004-04-30
: PRIOR APPLICATION NUMBER: PCT/IB02/04550
: PRIOR FILING DATE: 2002-10-31
: PRIOR APPLICATION NUMBER: ZA 2001/8978
: NUMBER OF SEQ ID NOS: 38
: SOFTWARE: PatentIn version 3.3
: SEQ ID NO 36
: LENGTH: 492
: TYPE: PRT
: ORGANISM: Human immunodeficiency virus type 1
US-10-494-131-36

Alignment Scores:

Pred. No.: 1.2e-118 Length: 492
Score: 2464.00 Matches: 466
Percent Similarity: 95.62% Conservative: 14
Best Local Similarity: 92.83% Mismatches: 12
Query Match: 86.76% Indels: 10
DB: 18 Gaps: 2

US-09-475-704A-4 (1-1509) x US-10-494-131-36 (1-492)

QY 1 ATGGGGCGCGCCAGCATCTCGCGCGGAGAGAGCTGGACAAGTGGAGAGAGATCCGC 60
Db 1 MetGlyAlaArgAlaSerIleLeuArgGlyGlyLysLeuAspLysTrpGlyLysIleArg 20
QY 61 CTGCGCGCGCGGCAAGAGACATACATGCTGAAGCACCTGTGTGGCCAGCGCGAG 120
Db 21 LeuArgProGlyGlyLysLysHisTyrMetLeuLysHisIleValTrpAlaSerArgGlu 40
QY 121 CTGAGAGGCTTGGCCCTGAACCCCGCGCTGTGAGAGACCGCGAGGCTGCAAGCAGATC 180
Db 41 LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluTrpSerGluGlyCysLysGlnIle 60
QY 181 ATGAAGAGCTGCAGCCCGCCCTGCAGACCGGACCGAGAGAGCTGCGAGCCTGTACAC 240
Db 61 MetLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluGluLeuLysSerLeuTyrAsn 80
QY 241 ACCGTGGCCACCTCTGACTGCTGTCAGCCCGGACATCGAGTCCGCGACACCAAGAGGCC 300
Db 81 ThrValAlaThrLeuTyrCysValHisGlyLysIleGluValArgAspThrLysGluAla 100
QY 301 CTGACACAGATCGAG 360
Db 101 LeuAspLysIleGluGluGluGlnAsnLysCysGlnGlnLysThrGlnGlnAlaLysAla 120
QY 361 GCCGAGCGCAAGGTGAGCCAGAACTACCCATCGTGCAGAACCTGCAGGGCCAGATGGTG 420
Db 121 AlaAspGlyLysValSerGlnAsnTyrProIleValGlnAsnLeuGlnGlyGlnMetVal 140
QY 421 CACCAAGCCATCAGCCCGCCGACCCCTGAACGCTGGGTGAAGGTGATGAGAGAGAGGCC 480
Db 141 HisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValIleGluGluLysAla 160
QY 481 TTCAGCCCGGAGGTGATCCCATGTTTCAACCGCCCTGAGCGAGGCGGCCACCCCGCAGAG 540
Db 161 PheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGlnAsp 180
QY 541 CTGAACACGATGTTGAACACCGTGGCGGCCACAGCGCCCATGACATGCTGAAGAGAC 600
Db 181 LeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLysAsp 200
QY 601 ACCATCAACGAGAGAGCGCGCGAGTGGACCGGCTGCACCCCGCTGCAGCGCGCGCCCTG 660
Db 201 ThrIleAsnGluGluAlaIleGluTrpAspArgValHisProValHisAlaGlyProIle 220
QY 661 GCGCGCGCGCAGATGCGCGAGACCCCGCGCGCAGCAGCATCGCGCGCGCCACAGACCTG 720
Db 221 AlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerThrLeu 240
QY 721 CAGAGACAGATCGCCTGGATGACCAAGAACCCCGCGTGCCTGGCGGAGACATCTAACAG 780
Db 241 GlnGluGlnIleAlaTrpMetThrSerAsnProProIleProValGlyAspIleTyrLys 260
QY 781 CGGTGATCATCTGGGCTTGAACAAGATCGTGGAGTGTACAGCCCGCTGAGCATCTG 840
Db 261 ArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProValSerIleLeu 280
QY 841 GACATCCGCCCAAGGCGCCCAAGAGCCCTTCCGCGACTACGTGAGCCGCTTCTCAAGACC 900
Db 281 AspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhePheLysThr 300
QY 901 CTGCGCGCGCAGAGAGGCCACCAAGAGAGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Db 301 LeuArgAlaGluGlnAlaThrGlnGluValLysAsnTrpMetThrAspThrLeuVal 320

QY 961 CAGAACGCCAACCCCGACTGCAGAACCATCTGCGCGCTCTCGGCGCCCGCCACCTG 1020
Db 321 GlnAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeuGlyProGlyAlaThrLeu 340
QY 1021 GAGAGATGATGACCGCCTGCCAGAGCGCTGGCGCGCCCGCCCAAGAGCCCGCTGCTG 1080
Db 341 GluGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLysValArgValLeu 360
QY 1081 GCCGAGCGGATGAGCCAGGCCCAAGCCGTGAACATCATGATGCAGAGAGCAACTTCAAG 1140
Db 361 AlaGluAlaMetSerGlnAlaAsnSerGlyAsnIleMetMetGlnArgSerAsnPheLys 380
QY 1141 GCGCGCGCGCGCAACGTCAGTGCCTTCAACTGCGGCAAGAGGGCCACATCGCCCAAGAC 1200
Db 381 GlyProArgArgIleValLysCysPheAsnCysGlyLysGluGlyHisIleAlaArgAsn 400
QY 1201 TGCGCGCGCGCGCGCAAGAGGGCTGCTGAAGTGCAGAGAGAGAGAGAGAGAGAGAG 1260
Db 401 CysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGluGlyHisGlnMetLys 420
QY 1261 GACTGACCCGAGCGCCCAAGGCCCACTTCTGGGCAAGATCTGGGCCACCAAGAGGCCCGC 1320
Db 421 AspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSerHisLysGlyArg 440
QY 1321 CCGCGCACTTCTGTCAGAACCCGAGCGAGCGCGCGCGCCCGCCACCGTGCACCGCCCGC 1380
Db 441 ProGlyAsnPheLeuGlnAsnArgProGlu-----ProThrAlaPro 454
QY 1381 CCGCGCGAGAGCTTCCGCTTGCAGAGAGACCAACCCCGCCCGCAAGAGAGCCCAAGAGAC 1440
Db 455 ProAlaGluSerPheArgPheGluGluThrThrProAlaProLysGlnGluProIleGlu 474
QY 1441 CGCGAGCCCTACCGCGAGCCCGCTGACCGCCCGCGCAGCCCTGTTCCGCGAGCGCGCCCTG 1500
Db 475 -----ArgGluProLeuThrSerLeuLysSerLeuPheGlySerAspProLeu 490
QY 1501 AGCCAG 1506
Db 491 SerGln 492

RESULT 5
US-09-991-258-5
; Sequence 5, Application US/09991258
; Patent No. US20020141975A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Robert
; APPLICANT: Keith, Paula
; APPLICANT: Dryga, Sergey
; APPLICANT: Caley, Ian
; APPLICANT: Maughan, Maureen
; APPLICANT: Johnston, Robert
; APPLICANT: Davis, Nancy
; APPLICANT: Swanson, Ronald
; TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOMES WITH MODIFIED HIV GENES FOR USE
; FILE REFERENCE: 01113.0001U3
; CURRENT APPLICATION NUMBER: US/09/991,258
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 09/902,537
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,995
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FaestSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 492
; TYPE: PRN
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. US20020141975A1e =
US-09-991-258-5

Alignment Scores:

Pred. No.:	2.74e-118	Length:	492
Score:	2457.00	Matches:	465
Percent Similarity:	95.22%	Conservative:	13
Best Local Similarity:	92.63%	Mismatches:	14
Query Match:	86.51%	Indels:	10
DB:	9	Gaps:	2

US-09-475-704A-4 (1-1509) x US-09-991-258-5 (1-492)

```

QY 1 ATGGCGCCGCGCCAGCATCTCGCGCGGAGAAAGCTGACAACTGGGAGAAGATCCGC 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 MetAlaAlaArgAlaSerIleLeuArgGlyGlyLeuAspLysTrpGlyLysIleArg 20

QY 61 CTGCGCCCGCGCGGCAAGAACACTATGCTGAAGCACTGTGTGGCCAGCCCGAG 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21 LeuArgProGlyGlyLysLysHisTyrMetLeuLysHisIleValTrpAlaSerArgL 40

QY 121 CTGAGGGCTTCGCTTGAACCCCGCTGTGAAGACCGCCGAGGGCTGCAAGAGATC 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 41 LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrSerGlyCysLysGlnIle 60

QY 181 ATGAAGCAGCTGCAGCCCGCTGTGACCGGACCGGAGAGAGCTGGCGAGCCTGTAC 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 MetLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluGluLeuLysSerLeuTyrAsn 80

QY 241 ACCGTGGCCACCTGTACTGCGTGCAGCGCGGATCGAGTCCGCGACCAAGAGAGCC 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 81 ThrValAlaThrLeuTyrCysValHisGluLysIleGluValArgAspThrLysGluAla 100

QY 301 CTGACAAGATCGAGAGAGAGAGACAAGTCCCGAGAGAGAGAGAGAGAGAGAGAG 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 101 LeuAspLysIleGluGluGluGlnAsnLysCysGlnGlnLysThrGlnGlnAlaLysAla 120

QY 361 GCCGACGGCAAGGTGAGCCAGAACTACCTCGTGCAGAACTGCAAGGCGCAGATG 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 AlaAspGlyLysValSerGlnAsnTyrProIleValGlnAsnLeuGlnGlyGlnMetVal 140

QY 421 CACCAAGCCATCAGCCCGCGCACCTTGAACGCTGGGTGAAGTGTATCGAGAGAGGCC 480
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 141 HisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValIleGluGluLysAla 160

QY 481 TTCAGCCCGGAGGTATCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCGCAGAC 540
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 161 PheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGlnAsp 180

QY 541 CTGAACAGATGTTGAACACCGTGGGGGCCACCAAGCCCGCATGACATGCTGAAGAC 600
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 LeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLysAsp 200

QY 601 ACCATCAACGAGAGAGCGCGCGAGTGGGACCGCTGCAACCCCGTGCAGGGCGCGCCGTG 660
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 201 ThrIleAsnGluGluAlaAlaGluTrpAspArgLysHisProValHisAlaGlyProIle 220

QY 661 GCCCGCGCGCAGATGGCGGACCCCGCGCGGAGCAGCATCGCCGCGCCACCGACCCCTG 720
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 221 AlaProGlyGlnMetCysArgGluProArgGlySerAspIleAlaGlyThrThrSerThrLeu 240

QY 721 CAGAGCAGATCGCTGATGACCAACCAACCCCGCTGCGCCCGTGGCGCAGATCTTACAAG 780
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 GlnGluGlnIleAlaTrpMetThrSerAsnProProIleProValGlyAspIleTyrLys 260

QY 781 CGGTGATCATCTGGGCTTGAACAGATCGTGGGATGTACAGCCCGCTGAGCATCTTG 840
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 261 ArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProValSerIleLeu 280

QY 841 GACATCCGCGCAGGGCGCCCAAGAGCCCTTCCGCGACTACGTGAGCCGCTTCTTCAAGACC 900
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 281 AspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhePheLysThr 300

QY 901 CTGCGCGCGGAGAGAGCCACCGAGAGCTGAAGACTGATGACCGAGACCCCTGCTGCTG 960
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 301 LeuArgAlaGluGlnAlaThrGlnGluValLysAsnTyrMetThrAspThrLeuLeuVal 320

QY 961 CAGAACGCCAACCCCGACTGCAAGAACCATCTCTCGCGGCTCTCGGCCCCCGCCACCCCTG 1020
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 321 GlnAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeuGlyProGlyAlaThrLeu 340

QY 1021 GAGAGATGATGACCCGCTGCGCAGGGCGGTGGCGGGCCCCCGCCACAAGSGCCGCTGCTG 1080
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 341 GluGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLysAlaArgValLeu 360

QY 1081 GCCGAGCGCATGAGCCAGGCCCAACAGCGTGACATCATGATGACAGAAGCAACTTCAAG 1140
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 AlaGluAlaMetSerGlnThrAsnSerGlyAsnIleMetMetGlnArgSerAsnPheLys 380

QY 1141 GGGCCCCGGCGCAACGTCAAGTCTTCAACTGCGGCAAGAGGGCCACATCGCCCAAGAAC 1200
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 381 GlyProArgArgIleValLysCysPheAsnCySGlyLysGluGlyHisIleAlaArgAsn 400

QY 1201 TGCCCGCGCCCCCGCAAGAGGGCTGCTGAACTGCGGCAAGAGGGCCACCATGAAG 1260
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 401 CysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGluGlyHisGlnMetLys 420

QY 1261 GACTGCACCGAGCGCCAGGCCCACTTCTGCGGCAAGATCTGCGCCACAGCAGAGGGCCGC 1320
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 421 AspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSerHisLysGlyArg 440

QY 1321 CCGGGCAACTTCTTGCAGAACCGCAGCGAGCGCGCGCCCAACCGTGCCCAACCGCCCC 1380
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 441 ProGlyAsnPheLeuGlnAsnArgProGlu-----ProThrAlaPro 454

QY 1381 CCGCGCGAGAGCTTCCGCTTGAAGAGAGACCAACCCCGCGCCCAAGCAGAGAGCCCAAGAC 1440
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 455 ProAlaGluSerPheArgPheGlnGluThrThrProAlaProLysGlnGluProIleGlu 474

QY 1441 CCGGAGCCCTACCGCGAGCGCCCTGACCGCCCTGCGCAGCCTGTTCGGCAGCGGCCCCCTG 1500
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 475 -----ArgGluProLeuThrSerLeuLysSerLeuPheGlySerAspProLeu 490

QY 1501 AGCCAG 1506
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 491 SerGln 492

RESULT 6
US-10-929-234-5
; Sequence 5, Application US/10929234
; Publication No. US2005012355SA1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Robert
; APPLICANT: Keith, Paula
; APPLICANT: Dryga, Sergey
; APPLICANT: Caley, Ian
; APPLICANT: Maughan, Maureen
; APPLICANT: Johnston, Robert
; APPLICANT: Davis, Nancy
; APPLICANT: Swansonstrom, Ronald
; TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOMES WITH MODIFIED HIV GENES FOR USE
; FILE REFERENCE: 01113.0001U3
; CURRENT APPLICATION NUMBER: US/10/929,234
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: US/09/991,258
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 09/902,537
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,995
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

```

OTHER INFORMATION: Description of Artificial Sequence; Note =
; OTHER INFORMATION: synthetic construct
US-10-929-234-5

Alignment Scores:

Pred. No.:	2.74e-118	Length:	492
Score:	2457.00	Matches:	465
Percent Similarity:	95.22%	Conservative:	13
Best Local Similarity:	92.63%	Mismatches:	14
Query Match:	86.51%	Indels:	10
DB:	18	Gaps:	2

US-09-475-704A-4 (1-1509) x US-10-929-234-5 (1-492)

```
QY      1 ATGGGCGCCGCGCCAGCATCTGCGCGGAGAGCTGGAACAAGTGGAGAGATCCGC 60
      11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MetAlaIaArGAlaSerIleLeuArgIglulysLeuAspIlystrpGluIlysIleArg 20

QY      61 CTGCGCGCCGCGCGCAAGAACACTATGCTGAAGCACTGTGTGGCCAGCGGCTGCAAGCAGATC 120
      61 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      21 LeuArgProgllygllysllyshistylMetleuIyshisilevaltrpAlaSerArgIlu 40

QY     121 CTGAGGGCTTGCCTTGAACCCCGGCTGTGAGACCGCGGAGGCTGCAAGCAGATC 180
      121 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      41 LeuGluArgPheAlaLeuAsnProgllyleuLeuGluThrSerGluIlyCysIysGlnIle 60

QY     181 ATGAAGCAGCTGCAGCCCGCTGCAGACCGGACCGGAGAGCTGCGACCTGTACAGC 240
      181 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      61 MetIysGlnLeuGlnProAlaLeuGlnThrGlyThrGluIleuIySerleuTyraSn 80

QY     241 ACCGTGGCCACCTGTACTGCGTGCAGACCGCGCATCGAGTCCGCCACACCAAGAGGCC 300
      241 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      81 ThrValAlaThrleuTyrcysvalHisGlnIlyslleGluValArgAspThrIySGluAla 100

QY     301 CTGACACAGATCGAGAGGAGAGACAGAACAGTCCCAAGAGAGACCAAGAGAGAGAG 360
      301 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      101 LeuAspIlyslleGluIglulGlnAsnIlyscysGlnIlyThrGlnIlyAlaIlyAla 120

QY     361 GCCGAGCGCAAGGTGAGCCAGACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGTTG 420
      361 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      121 AlaAspGlyIlyvalSerGlnAsnTyProIleValGlnAsnLeuGlnIlyGlnMetVal 140

QY     421 CACCAAGCCATCAGCCCGCCGACCCCTGAACGCTGGGTGAAGGTGATGAGAGAGAGCC 480
      421 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      141 HisGlnAlaIleSerProArgThrLeuAsnAlaIleTyrValIlyvalIleGluIlyAla 160

QY     481 TTCAGCCCGGAGGTGATCCCATGTTTCAACCGCCCTGAGCGAGGGCGGCCACCCCGCAG 540
      481 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      161 PheSerProGluValIleProMetPheThrAlaLeuSerGluIlyAlaThrProGlnAsp 180

QY     541 CTGAACACGATGTTGAACACCGTGGGGCGGCCACCAAGCGCCATGACATGCTGAAGGAC 600
      541 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      181 LeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuIyAsp 200

QY     601 ACCATCAACGAGAGGGCGCCGAGTGGGACCGCTGCACCCCGTGCAGGGCCGGCCCGCTG 660
      601 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      201 ThrIleAsnGluGluAlaIaGluIlyrAspArgLeuHisProValHisAlaGlyProIle 220

QY     661 GCCCGCGCCAGATGCGCGGACCCCGCGGCGAGCAGCATCGCGCGGCCACCAAGCAGCCCTG 720
      661 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      221 AlaProgllyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerThrLeu 240

QY     721 CAGGAGCAGATCGCTGTGATGACCAAGACCCCGCTGCGCGGCGGACATCTTACAAG 780
      721 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      241 GlnGluGlnIleAlaIleTyrMetThrSerAsnProIleProValGlyAspIleIyLys 260

QY     781 CGGTGATCATCTTGGGCTGAACAAGATCGTGGATGTACAGCCCGCTGAGCATCTG 840
      781 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      261 ArgTyrIleIleLeuGlyLeuAsnIlyIleValArgMetTySerProValSerIleLeu 280

QY     841 GACATCCGCGCAGGGCGCCCAAGAGAGCCCTTCCGCGACACTACGTGAGCCGCTTCTTCAAG 900
      841 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      281 AspIleArgGlnGlyProIySGluProPheArgAspTyValAspArgPhePheIyThr 300
```

```
QY      901 CTGCGCGCCGAGAGGCCACCCAGAGAGCTGAAGACTGATGACCGAGACCTGTGCTG 960
      901 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      301 LeuArgAlaGluGlnAlaThrGlnGluValIlyAsnTrpMetThrAspThrLeuVal 320

QY      961 CAGAACGCCAACCCCGACTGCAGAACCATCTGCGCGCTCTCGGCCCCGGGCCACCCCTG 1020
      961 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      321 GlnAsnAlaAsnProAspCysIystrIleLeuArgAlaLeuGlyProGlyAlaThrIleu 340

QY     1021 GAGAGATGATGACCGCTGCGGAGGGCGGTGGCGGCCCGGCCACAGGCCCGCTGCTG 1080
      1021 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      341 GluIleMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisIlyslAlaArgValIleu 360

QY     1081 GCCGAGCGGATGAGCCCAAGGCCCAACAGCGTGAACATCATGTGCAAGAGCAACTTCAAG 1140
      1081 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      361 AlaGluAlaMetSerGlnThrAsnSerGlyAsnIleMetMetGlnArgSerAsnDhels 380

QY     1141 GCGCGCGCGCGCAACGTCAAGTGTCTTCAACTGCGGCAAGAGGGCCACATCGCCCAAGAAC 1200
      1141 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      381 GlyProArgArgIleValIlyscysPheAsnCyGlyIySGluGlyHisIleAlaArgAsn 400

QY     1201 TGCGCGCGCGCGCGCAAGAGGGCTGTGAGATGTCGCGCAAGAGGGCCACAGATGAAG 1260
      1201 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      401 CysArgAlaProArgIlysllySGlyCysTrpIyCysGlyIySGluGlyHisGlnMetIys 420

QY     1261 GACTGCACCGAGCGCCAGGCCAACTTCTGGGCAAGATCTGGCCCAAGCAGAGGGCCGC 1320
      1261 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      421 AspCysThrGluArgGlnAlaAsnPheLeuGlyIlyslletrProSerHisIySGlyArg 440

QY     1321 CCCGCAACTTCTTCAGAACCCGAGCGGCGCGCGCCCGCCACCGTGCACCGCCCGCC 1380
      1321 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      441 ProGlyAsnPheLeuGlnAsnArgProGlu-----ProThrAlaPro 454

QY     1381 CCCGCGAGAGCTTCCGCTTGAAGAGACCAACCCCGCCCGCCCAAGAGAGCCCAAGGAC 1440
      1381 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      455 ProAlaGluSerPheArgPheGluGluThrThrProAlaProIySGlnGluProIleGlu 474

QY     1441 CCGGAGCCCTACCGCGAGCCCGCTGACCGCCCTGCGCAGCCTGTTGGCAGCGGCCCTG 1500
      1441 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      475 -----ArgGluProLeuThrSerLeuIySerLeuPheGlySerAspProLeu 490

QY     1501 AGCCAG 1506
      1501 | | | | |
Db      491 SerGln 492

RESULT 7
US-10-332-413-8
; Sequence 8, Application US/10332413
; Publication No. US2004011660A1
; GENERAL INFORMATION:
; APPLICANT: Johnstone, Robert Edward
; APPLICANT: Swanstrom, Ronald Ivar
; APPLICANT: Morris, Lynn
; APPLICANT: Karim, Salim Abdool
; APPLICANT: Williamson, Carolyn
; TITLE OF INVENTION: Process for the Selection of HIV-1 Subtype C Isolates, Selected H
; FILE REFERENCE: 45669-281993
; CURRENT APPLICATION NUMBER: US/10/332,413
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: PCT/IB01/01208
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,995
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: ZA 2000/3437
; PRIOR FILING DATE: 2000-07-10
; PRIOR APPLICATION NUMBER: ZA 2000/4924
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 631
; TYPE: PRT
```

```
; ORGANISM: Human immunodeficiency virus type 1
US-10-332-413-8

Alignment Scores:
Pred. No.:      2.76e-118      Length:      631
Score:          2457.00        Matches:     465
Percent Similarity: 95.22%      Conservative: 13
Best Local Similarity: 92.63%      Mismatches:  14
Query Match:    86.51%         Indels:      10
DB:             16             Gaps:        2

US-09-475-704A-4 (1-1509) x US-10-332-413-8 (1-631)

QY      1 ATGGGCGCCGCGCCAGCATCTGCGGCGGAGAAAGCTGACAAGTGGAGAGAAGATCCGC 60
Db      31 MetalaAlaArgAlaSerileLeuArgGlyGluLeuAspLysTrpGluLysileArg 50
QY      61 CTGCGCCCGCGCGGCAAGAACACTACATGCTGAAGCACTGTGTGGCCAGCCGCGAG 120
Db      51 LeuArgProGlyGlyLysLysHisTrpMetLeuLysHisileValTrpAlaSerArgGlu 70
QY      121 CTGAGGGGCTTCGCGCTGTAACCCCGGCTGCTGAAGACCGCGAGGCTGCAAGCAGATC 180
Db      71 LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrSerGluGlyCysLysGlnile 90
QY      181 ATGAAGCAGCTGCAGCGCCCGCTGACAGACCGGACCGAGAGCTGGCGCAGCCTGTACAAC 240
Db      91 MetLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluGluLeuLysSerLeuTrpAsn 110
QY      241 ACCGTGGCCACCGCTGTACTGCGTGACCGCGCATGAGGTCCGCGACACCAAGAGGCC 300
Db      111 ThrValAlaThrLeuTrpCysValHisGluLysileGluValArgAspThrLysGluAla 130
QY      301 CTGACAAGATCGAGAGGAGAGCAGAAACAAGTCCACAGAGAAGACCAAGCCAGAGAG 360
Db      131 LeuAspLysileGluGluGluGlnAsnLysCysGlnGlnLysThrGlnGlnAlaLysAla 150
QY      361 GCCGACGGCAAGGTGAGCCAGAACTACCCCATCGTGCAAGAACCTGCAAGGCGCAGATG 420
Db      151 AlaAspGlyLysValSerGlnAsnTrpProileValGlnAsnLeuGlnGlyGlnMetVal 170
QY      421 CACCAAGCCATCAGCGCCCGCGACACCTGAAGCGCTGGTGAAGTGAAGTGAAGAGGCC 480
Db      171 HisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValileGluGluLysAla 190
QY      481 TTCAGCCCCGAGGTGATCCCATGTTTACCGCCCTGAGCGAGGCGCCACCCCGCAGAG 540
Db      191 PheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGlnAsp 210
QY      541 CTGAACACGATGTTGAACAACCGTGGGGGCCACCAAGCCCGCATGACATGCTGAAGAC 600
Db      211 LeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLysAsp 230
QY      601 ACCATCAACGAGAGGCGCGCGAGTGGGACCGCCTGCAACCCCGTGACGCGCGCCCGTG 660
Db      231 ThrIleAsnGluGluAlaAlaGluTrpAspArgLysHisProValHisAlaGlyProile 250
QY      661 GCGCGCGCGCAGATGCGCGCGACCCCGCGGCGACGCAATCGCGCGCGCCACCAAGCCTG 720
Db      251 AlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerThrLeu 270
QY      721 CAGAGCAGATCGCGCTGATGACCAAGCAACCCCGCTGCCCGTGGCGGACATCTACAAG 780
Db      271 GlnGluGlnIleAlaTrpMetThrSerAsnProProileProValGlyAspIleTrpLys 290
QY      781 CGGTGATCATCTGGGCTTGAACAAGATCGTGGGATGTACAGCCCGCTGAGCATCTG 840
Db      291 ArgTrpIleileLeuGlyLeuAsnLysileValArgMetLysSerProValSerileLeu 310
QY      841 GACATCCGCGCAGGCGCCCAAGAGCGCTTCCGCGACTACGTGAGCGGCTTCTTAAGACC 900
Db      311 AspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhePheLysThr 330
```

```
QY      901 CTGCGCGCCGAGAGGCCCAACCCAGAGACGTGAAGAAGTGAATGACCGAGACCTGTGCTG 960
Db      331 LeuArgAlaGluGlnAlaThrGlnGluValLysAsnTrpMetThrAspThrLeuLeuVal 350
QY      961 CAGAACGCCCAACCCCGACTGCAAGACCATCTCTGCGCGCTCTGCGCGCCCGCCACCTG 1020
Db      351 GlnAsnAlaAsnProAspCysLysThrileLeuArgAlaLeuGlyProGlyAlaThrLeu 370
QY      1021 GAGAGATGATGACCGCGCTGCCAGGGCGGTGGCGCGCGCCCAAGAGCGCGCTGCTG 1080
Db      371 GluGluMetMetThrAlaCysGlnGlnLysValGlyGlyProGlyHisLysAlaArgValLeu 390
QY      1081 GCCGAGCGGATGACGCCCAAGGCCCAACAGCGTGAACATCATGATGCAAGAAGCAACTTCAAG 1140
Db      391 AlaGluAlaMetSerGlnThrAsnSerGlyAsnIleMetMetGlnArgSerAsnPheLys 410
QY      1141 GCGCGCCCGCGCAACGTCAAGTGTCTTCACTGCGCGCAAGAGGGCCACATCGCCCAAGAAC 1200
Db      411 GlyProArgArgIleValLysCysPheAsnCysGlyLysGluGlyHisileAlaArgAsn 430
QY      1201 TGCGCGCGCGCGCGCAAGAGGGCTGCTGAAGTGGGCAAGAGGGCCACCAAGATGAAG 1260
Db      431 CysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGluGlyHisileGlnMetLys 450
QY      1261 GACTGCACCGAGCGCCAGGCGCAACTTCTGCGCAAGATCTGCGCCACCAAGAGGGCGCG 1320
Db      451 AspCysThrGluArgGlnAlaAsnPheLeuGlyLysileTrpProSerHisLysGlyArg 470
QY      1321 CCGGCGCACTTCTGCAAGACCGCAGCGAGCGCGCGCGCCCAACGCTGCCACCGCGCCC 1380
Db      471 ProGlyAsnPheLeuGlnAsnArgProGlu-----ProThrAlaPro 484
QY      1381 CCGCGCGAGAGCTTCGCTTGCAGAGAGACCAACCCCGCGCCCAAGCAGAGGCCCAAGGAC 1440
Db      485 ProAlaGluSerPheArgPheGluGluThrThrProAlaProLysGlnGluProileGlu 504
QY      1441 CCGGAGCGCTTACCGCGAGCGCGCTGACCGCGCTGCGCAGCCTGTTCGGCAGCGCGCTG 1500
Db      505 -----ArgGluProLeuThrSerLeuLysSerLeuPheGlySerAspProLeu 520
QY      1501 AGCCAG 1506
Db      521 SerGln 522

RESULT 8
US-10-780-507-98
; Sequence 98, Application US/10780507
; Publication No. US20050137387A1
; GENERAL INFORMATION:
; APPLICANT: MULLINS, James I.
; APPLICANT: RODRIGO, Allen G.
; APPLICANT: LEARN, Gerald H.
; APPLICANT: LI, Fusheng
; APPLICANT: NICKLE, David C.
; APPLICANT: JENSEN, Mark A.
; TITLE OF INVENTION: ANCESTRAL AND COT VIRAL SEQUENCES, PROTEINS AND IMMUNOGENIC COMPO
; FILE REFERENCE: 16336-001320US
; CURRENT APPLICATION NUMBER: US/10/780,507
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US 10/204,204
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US 60/447,586
; PRIOR FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 98
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Artificial sequence
```


FEATURE:
; OTHER INFORMATION: Minimum of means center of tree reconstruction of clade C gag prc
; OTHER INFORMATION: tein sequence
US-10-780-507-98

Alignment Scores:

Pred. No.:	2.91e-118	Length:	493
Score:	2456.50	Matches:	468
Percent Similarity:	95.43%	Conservative:	12
Best Local Similarity:	93.04%	Mismatches:	12
Query Match:	86.50%	Indels:	11
DB:	18	Gaps:	3

US-09-475-704A-4 (1-1509) x US-10-780-507-98 (1-493)

```
QY 1 ATGGGCGCCGCGGCGCATCTCGCGGCGGAGAGCTGAGCAAGTGGGAGAAGATCCGC 60
    |||||||
Db 1 MetGlyAlaArgAlaSerIleLeuArgGlyGlyLysLeuAspThrTrpGluLysIleArg 20

QY 61 CTGGCGCCCGCGGCGGAGAGCACTAATGCTGAAGCACTGTGTGGGCCAGCCGCGAG 120
    |||||||
Db 21 LeuArgProGlyGlyLysLysHisTyrMetLeuLysHisLeuValTrpAlaSerArgGlu 40

QY 121 CTGAGGGGCTTCGCCCTGAACCCCGGCTGTGTGAGACCGCGGAGGCTGCAAGCATC 180
    |||||||
Db 41 LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrSerGluGlyCysLysGlnIle 60

QY 181 ATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGACCGAGAGCTGCCAGCCTGTACAAC 240
    |||||||
Db 61 MetLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluGluLeuArgSerLeuTyrAsn 80

QY 241 ACCGTGCGCCACCCCTGTACTGCGTGCACGCGCGGATGAGGTCCGCCACCAAGAGGCC 300
    |||||||
Db 81 ThrValAlaThrLeuTyrCysValHisGluLysIleGluValArgAspThrLysGluAla 100

QY 301 CTGACAAAGATCGAGAGAGAGCAAGAAAGTCCCGACAGAGAGACCCAGAGGCCAAG--- 357
    |||||||
Db 101 LeuAspLysIleGluGluGluGlnAsnLysSerGlnGlnLysThrGlnGlnAlaGluAla 120

QY 358 GAGGCGGAGGCGAAGGTGAGCCAGAACTACCCCATGTGCAAGACCTGCAGGGCGCAGATG 417
    |||||||
Db 121 AlaAlaAspGlyLysValSerGlnAsnTyrProIleValGlnAsnLeuGlnGlyMet 140

QY 418 GTGCACCAAGGCCATCAGCCCCCGGACCTGAAACGCTGGGTGAAGGTGATGAGAGAAAG 477
    |||||||
Db 141 ValHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysValIleGluGluLys 160

QY 478 GCGTTTCAGCCCGGAGGTGATCCCATGTTCAACGCGCTGAGCGGAGGCGCCACCCCGCAG 537
    |||||||
Db 161 AlaPheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGln 180

QY 538 GACCTGAACACGATGTTGAACACACCGTGGGCGGCGCACAGCGCGCATGACAGATGCTGAAG 597
    |||||||
Db 181 AspLeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLys 200

QY 598 GACACCATCAACGAGAGAGGCGCGCGAGTGGACCGCTGCACCCCGTGCAGGCGCGCCCC 657
    |||||||
Db 201 AspThrIleAsnGluGluAlaAlaGluTrpAspArgLeuHisProValHisAlaGlyPro 220

QY 658 GTGGCCCCCGGCGCAGATGCGGAGCCCGCGGCGGCGCAGGACATCGCCGCGCCACCGCACCC 717
    |||||||
Db 221 ValAlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerThr 240

QY 718 CTGCAGGAGCAGATGCTGCTGATGACCAAGCAACCCCGCTGCGCGTGGGCGCACTTAC 777
    |||||||
Db 241 LeuGlnGluGlnIleAlaTrpMetThrSerAsnProProValGlyAspIleTyr 260

QY 778 AAGCGGTGATCATCTGGGCTGGAACAAAGATCGTGGGATGTACAGCCCGTGCATC 837
    |||||||
Db 261 LysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProValSerIle 280

QY 838 CTGACATCCGCGCAGGCGCCCAAGAGCCCTTCGCGACTACGTGACCGCTTCTTCAAG 897
    |||||||
```

```
Db 281 LeuAspIleLysGlnGlyProLysGluProPheArgAspTyrValAspArgPhePheLys 300
QY 898 ACCCTGCGCGCGGAGCAGAGGCCACCCAGAGCATGAAGAACTGGATGACCGACCTGCTG 957
    |||||||
Db 301 ThrLeuArgAlaGluGlnAlaThrGlnAspValLysAsnTrpMetThrAspThrLeuLeu 320

QY 958 GTGCAGAAAGCCAAACCCCGACTGCAAGACCATCTGCGGCTCTGGCCCCGCGCCACC 1017
    |||||||
Db 321 ValGlnAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeuGlyProGlyAlaThr 340

QY 1018 CTGAGAGAGATGATGACCGCTGCGCAGGCGCGTGGGCGGCGCCACAAAGCGCGTGT 1077
    |||||||
Db 341 LeuGluGluMetMetThrAlaCysGlnGlyValGlyLysProGlyHisLysAlaArgVal 360

QY 1078 CTGGCCGAGGCGCATGAGCCAGGCGCAACGCTGAACATCATGATGACAGAGCACTTC 1137
    |||||||
Db 361 LeuAlaGluAlaMetSerGlnAlaAsnAsnThrAsnIleMetMetGlnArgSerAsnPhe 380

QY 1138 AAGGCGCCCGCGGCGCAACGTCAAGTCTTCAACTGCGGCAAGAGGCGCCACATCGCCAAG 1197
    |||||||
Db 381 LysGlyProLysArgIleValLysCysPheAsnCysGlyLysGluGlyHisIleAlaArg 400

QY 1198 AACTGCGCGCGCCCGCGCAAGAGGGCTGTGTAAGTGCAGGAGAGGCGCCACAGATG 1257
    |||||||
Db 401 AsnCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGluGlyHisGlnMet 420

QY 1258 AAGGACTGCACCGGCGGCGGCAACTTCTGGGCAAGATCTGGCCAGCCCAAGGGC 1317
    |||||||
Db 421 LysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSerHisLysGly 440

QY 1318 CGCCCCGCAACTTCTTCAGAACCGGACGAGCGCCGCGCCCAACCGTGCACCGGCC 1377
    |||||||
Db 441 ArgProGlyAsnPheLeuGlnSerArgProGlu-----ProThrAla 454

QY 1378 CCCCCCGCGAGAGCTTCCGCTTCGAGAGAGACACCCCGCCCAAGCAGAGCCCAAG 1437
    |||||||
Db 455 ProProAlaGluSerPheArgPheGluGluThrThrProAlaProLysGlnGluProLys 474

QY 1438 GACCGCGAGCCCTACCGGAGCGCCCTGACCGCGCTGCGCAGCTGTTCGCGAGCGGCC 1497
    |||
Db 475 Asp-----ArgGluProLeuThrSerLeuLysSerLeuPheGlySerAspPro 490

QY 1498 CTGAGCCAG 1506
    |||||||
Db 491 LeuSerGln 493

RESULT 9
US-10-780-507-96
; Sequence 96, Application US/10780507
; Publication No. US20050137387A1
; GENERAL INFORMATION:
; APPLICANT: MULLINS, James I.
; APPLICANT: RODRIGO, Allen G.
; APPLICANT: LEARN, Gerald H.
; APPLICANT: LI, Fusheng
; APPLICANT: NICKLE, David C.
; APPLICANT: JENSEN, Mark A.
; TITLE OF INVENTION: ANCESTRAL AND COT VIRAL SEQUENCES, PROTEINS AND IMMUNOGENIC COMPO
; FILE REFERENCE: 16336-001320US
; CURRENT APPLICATION NUMBER: US/10/780,507
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US 10/204,204
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: PCT/US01/05288
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/183,659
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: US 60/447,586
; PRIOR FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 96
; LENGTH: 492
```

```

; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Most recent common ancestor reconstruction of clade C gag protein
; OTHER INFORMATION: sequence
US-10-780-507-96

```

Alignment Scores:

Pred. No.:	6.28e-118	Length:	492
Score:	2450.00	Matches:	463
Percent Similarity:	95.22%	Conservative:	15
Best Local Similarity:	92.23%	Mismatches:	14
Query Match:	86.27%	Indels:	10
DB:	18	Gaps:	2

US-09-475-704A-4 (1-1509) x US-10-780-507-96 (1-492)

```

QY 1 ATGGGCGCCGCGCAGCATCTGCGCGCGAGAACTGGACAAGTGGAGAGATCCGC 60
Db 1 MetGlyAlaArgAlaSerIleLeuArgGlyGlyLeuAspThrTrpGlyIleArg 20
QY 61 CTGCGCCCGCGCGCAGAAAGCACTACATGCTGAAGCACTGTGTGGCCAGCCGCGAG 120
Db 21 LeuArgProGlyGlyblybshIstYrMetIleYshIbLeuValTrpAlaSerArgGlu 40
QY 121 CTGAGGGCTTCGCGCTGAACCCCGCGCTGCTGAAGCCGCGAGGCTGCAAGCATC 180
Db 41 LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrSerGluGlyCysIleGlnIle 60
QY 181 ATGAGCAGCTGCAGCGCCGCGCTGCAAGCCGCGCAGGAGAGCTGCGCACTGTACAAC 240
Db 61 IleYsgInLeuGlnProAlaLeuGlnThrGlyThrGluGluLeuYsSerLeuYrAsn 80
QY 241 ACCGTGGCCACCTCTACTGCTGCTGCAAGCCGCGCAGGAGGTCCGCGCAACCAAGAGGCC 300
Db 81 ThrValAlaThrLeuYrCysValHisGlnArgIleGluValArgAspThrIleGlnAla 100
QY 301 CTGCAACAAGATCGAGAGAGAGAGACAAGTCCCGCAGAGAGAGAGAGAGAGAGAGAG 360
Db 101 LeuAspIlyIleGluGluGluGlnAsnIlySerGlnGlnIlyThrGlnGlnAlaGluAla 120
QY 361 GCCGACGGCAAGGTGAGCCAGAACTACCCCATCGCTGTCAGAACCTGCAAGGCGCAGATGTTG 420
Db 121 AlaAspGlyIlyValSerGlnAsnYrProIleValGlnAsnLeuGlnIlyGlnMetVal 140
QY 421 CACCAAGGCATCAGCCCCCGCACCTGAAAGCGCTGGGTGAAGTGAAGTGAAGAGAGGCC 480
Db 141 HisGlnAlaIleSerProArgThrLeuAsnAlaIleTrpValIlyValIleGluGluIlybAla 160
QY 481 TTCAGCCCCGAGGTGATCCCATGTTTCAACCGCCCTGAGCGAGGCGGCCACCCCGCAGGAC 540
Db 161 PheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGlnAsp 180
QY 541 CTGAACACGATGTTGAACACCGTGGCGGCCACACGCGCCCATGACATGCTGAAGGAC 600
Db 181 LeuAsnThrMetLeuAsnThrValGlyGlnHisGlnAlaAlaMetGlnMetLeuIlyAsp 200
QY 601 ACCATCAACGAGAGGCGCGCGAGTGGAGACCGCTGCAACCCCGTGCAGGCGCGCCCGTGG 660
Db 201 ThrIleAsnGluGluAlaAlaGluTrpAspArgLeuHisProValHisAlaGlyProVal 220
QY 661 GCCCGCGCGCAGATGCGCGACCCCGCGCGCAGCAGACATGCGCGCGGCCACACGACCCCTG 720
Db 221 AlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerThrLeu 240
QY 721 CAGAGCAGATCGCTGATGACACGACCAACCCCGCTGCCCGCGCAGCATCTACAAG 780
Db 241 GlnGluGlnIleAlaIleTrpMetThrSerAsnProProIleProValGlyAspIleYrIly 260
QY 781 CGGTGATCATCTCGGCGCTGAACAAGATCGTGGAGTGTACAGCCCGCTGAGCATCTCG 840
Db 261 ArgTrpIleIleLeuGlyLeuAsnIlyIleValArgMetYrSerProValSerIleLeu 280

```

```

QY 841 GACATCCGCCAGGGCCCCCAAGAGAGCCCTTCCGCGACTACGTGAGACCGCTTCTTCAAGACC 900
Db 281 AspIleYsgInGlyProIlyGluProPheArgAspTyrValAspArgPhehelYsThr 300
QY 901 CTGCGCGCGCGCAGAGGCCACCCAGAGACGCTGAAGAACTGATGACCGAGACCTGTGCTG 960
Db 301 LeuArgAlaGluGlnAlaIleThrGlnAspValIlybAsnTrpMetThrAspThrLeuVal 320
QY 961 CAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCTCTGCGGCCCGCGCCACCCCTG 1020
Db 321 GlnAsnAlaAsnProAspCysIlyThrIleLeuArgAlaLeuGlyProGlyAlaThrLeu 340
QY 1021 GAGAGATGATGACCGCGCTGCGAGGGCGCTGGCGCGGCCCGCCACAGGCGCGCTGCTG 1080
Db 341 GluGluMetMetThrAlaCysGlnGlyValGlyGlyProSerHisIlybAlaArgValLeu 360
QY 1081 GCCGAGCGCATGACCCAGGCCCAACAGCGTGAACATCATGATGCAAGAGCAACTTCAAG 1140
Db 361 AlaGluAlaMetSerGlnAlaAsnAsnThrAsnIleMetMetGlnArgGlyAsnPhelys 380
QY 1141 GCGCCCGCGCGCAACGTCAAGTCTTCAACTGCGCGCAAGAGGGCCACATCGCCCAAGAAC 1200
Db 381 GlyProArgArgIleValIlyCysPheAsnCybGlyIlybGluGlyHisIleAlaArgAsn 400
QY 1201 TGCCCGCGCGCGCGCAGAGAGGCGCTGCTGAAGTCCGCGCAAGAGGGCCACCATGAAG 1260
Db 401 CysArgAlaProArgIlybIlyGlyCysTrpIlybCysGlyIlybGluGlyHisGlnMetIly 420
QY 1261 GACTGCACCGAGCGCCAGGCCCACTTCTGCGCAAGATCTGCGCCCAAGAGAGGGCGCG 1320
Db 421 AspCysThrGluArgGlnAlaAsnPhelGluGlyIleTrpProSerHisIlybGlyArg 440
QY 1321 CCGCGCACTTCTTCCAGAACCGCAGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Db 441 ProGlyAsnPhelGluGlnSerArgProGlu-----ProThrAlaPro 454
QY 1381 CCGCGCGAGAGCTTCGCTTCCAGAGAGAGACCAACCCCGCGCGCGCGCGCGCGCGCGCG 1440
Db 455 ProAlaGluSerPheArgPheGluGluThrThrProAlaProIlybGlnGluProIlyAsp 474
QY 1441 CGCGAGCCCTACCGCGAGCGCGCTGACCGCGCTGCGCAGCGCTGTTGCGCAGCGCGCGCTG 1500
Db 475 -----ArgGluProLeuThrSerLeuIlybSerLeuPheGlySerAspProLeu 490
QY 1501 AGCCAG 1506
Db 491 SerGln 492

```

RESULT 10

US-10-339-217-109

; Sequence 109, Application US/10339217

; Publication No. US20030198648A1

; GENERAL INFORMATION:

; APPLICANT: Buechter, Douglas

; APPLICANT: Hou, Xiaohong

; APPLICANT: Marlor, Christopher W.

; APPLICANT: Rice, William G.

; TITLE OF INVENTION: Methods for Identifying Compounds Which Inhibit Binding of

; TITLE OF INVENTION: Nucleocapsid 7 Protein to HIV-1 RNA

; FILE REFERENCE: 111021.143(ACH-US1)

; CURRENT APPLICATION NUMBER: US/10/339,217

; CURRENT FILING DATE: 2003-01-09

; PRIOR APPLICATION NUMBER: US 60/347,369

; PRIOR FILING DATE: 2002-01-11

; NUMBER OF SEQ ID NOS: 144

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 109

; LENGTH: 492

; TYPE: PRT

; ORGANISM: HIV-1

US-10-339-217-109

Alignment Scores:

Pred. No.: 2.05e-117 Length: 492
Score: 2440.00 Matches: 464
Percent Similarity: 94.42% Conservative: 10
Best Local Similarity: 92.43% Mismatches: 18
Query Match: 85.92% Indels: 10
DB: 14 Gaps: 2

US-09-475-704A-4 (1-1509) x US-10-339-217-109 (1-492)

QY 1 ATGGGGCCCCGGCCAGCATCTTCGGGGCGAGAGCTGGACAAGTGGGAGAAGATCCGC 60
Db 1 MetGlyAlaSerAlaSerIleLeuArgGlyGlyLeuAspLysTrpGluLysIleArg 20
QY 61 CTGCGCCCCGGCGGCAAGACACTACATGCTGAAGCACTGTGTGGGCCAGCCGCGAG 120
Db 21 LeuArgProGlyGlyLysLysLysTrpArgLeuLysHisLeuValTrpAlaSerArgGlu 40
QY 121 CTGAGGGCTTCGCTGAAACCCCGGCTGCTGAGACCGCCGAGGGCTGCAAGCAGATC 180
Db 41 LeuGluArgPheAlaLeuAsnSerGlyLeuLeuGluThrAlaGluGlyCysLysGlnIle 60
QY 181 ATGAGCAGCTGCAGCCCGCCCTGCAGACCGGACGAGAGCTGCGCACCTGTACAAC 240
Db 61 IleLysGlnLeuGlnProAlaLeuGlnThrGlyThrGluGluLeuLysSerLeuTrpAsn 80
QY 241 ACCGTGGCCACCTGTACTGCGTGCACGCCGACATCGAGGTCCGCGACACCAAGAGGCC 300
Db 81 ThrValAlaThrLeuTrpCysValHisAlaGlyIleGluValArgAspThrLysGluAla 100
QY 301 CTGCAACAAGATCGAGAGAGAGAGACAAGTCCCGAGAGAGAGAGAGAGAGAGAGAG 360
Db 101 LeuAspLysIleGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 120
QY 361 GCCGACGGCAAGGTGAGCCAGAACCTACCTGTCAGAACCTGACGGCCAGATGGTG 420
Db 121 AlaAspGlyLysValSerGlnAsnTrpProIleValGlnAsnLeuGlnGlyGlnMetVal 140
QY 421 CACCAAGCCATCAGCCCCCGCACCTGTAACGCTGGGTGAAGTGAATCGAGAGAGAGGCC 480
Db 141 HisGlnProIleSerProArgThrLeuAsnAlaTrpValLysValIleGluGluLysAla 160
QY 481 TTCAAGCCCGAGGTGATCCCATGTTCAACCGCCCTGAGCGAGGGCCGACCCCGCAGAC 540
Db 161 PheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGlnAsp 180
QY 541 CTGAACAAGATGTTGAACACCGTGGGGGGCCACCAAGCCCGCATGACATGCTGAAGAC 600
Db 181 LeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLysAsp 200
QY 601 ACCATCAACGAGAGAGCGCCCGAGTGGGAGACCGCTGCAACCCCGTGCAGCGCCCGCTG 660
Db 201 ThrIleAsnGluGluAlaAlaGluTrpAspArgLeuHisProValHisAlaGlyProVal 220
QY 661 GCCCGCCCGCAGATGCGCGCAGACCCCGGGGAGAGAGATCGCCGCGCCACCAAGACCTG 720
Db 221 AlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerThrLeu 240
QY 721 CAGAGCAGATCGCTGATGACCAAGACCCCGCTGCGCGCTGGGGCGACATCTACAG 780
Db 241 GlnGluGlnIleThrTrpMetThrSerAsnProProValProValGlyAspIleTrpLys 260
QY 781 CGGTGATCATCTTCGGCTTGAACAAGATCGTGGGATGTACAGCCCGCTGAGCATCTG 840
Db 261 ArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTrpSerProValSerIleLeu 280
QY 841 GACATCGCCAGAGGGCCCAAGAGAGCCCTTCGCGACTACGTGACCCGCTTCTCAAGACC 900
Db 281 AspIleLysGlnGlyProLysGluProPheArgAspTrpValAspArgPhePheLysVal 300
QY 901 CTGCGCGCCGAGAGAGGCCACCAAGAGAGTGAAGAACTGGATGACCGAGACCTGCTGGTG 960
Db 301 LeuArgAlaGluGlnAlaThrGlnAspValLysAsnTrpMetThrAspThrLeuLeuVal 320

QY 961 CAGACGCCAACCCCGACTGCAAGACCATCTGCGGCTCTGCGCCCGCGCCACCCCTG 1020
Db 321 GlnAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeuGlyProAlaAlaSerLeu 340
QY 1021 GAGGAGATGATGACCGCTGCGAGGGCGGTGGGGCGCCCGCCACCAAGGCGCGTGTG 1080
Db 341 GluGluMetMetThrAlaCysGlnGlyValGlyGlyProSerHisLysAlaArgValLeu 360
QY 1081 GCCGAGCGGATGAGCCAGCGCCACAGCGTGAAATCATGATGACGAAGACAACTTCAAG 1140
Db 361 AlaGluAlaMetSerGlnAlaAsnThrThrAsnIleMetMetGlnLysSerAsnPheLys 380
QY 1141 GGCCCCCGCGCAACGTCAAGTCTTCACTGCGGCAAGAGGGCCACATCGCCCAAGAAC 1200
Db 381 GlyProArgArgThrValLysCysPheAsnCysGlyLysGluGlyHisIleAlaLysAsn 400
QY 1201 TGCCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGAGGGCCACCATGTAAG 1260
Db 401 CysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGluGlyHisGlnMetLys 420
QY 1261 GACTGCACCGAGCGCCAGGCCCACTTCTGCGCAAGATCTGGGCCACGCCACAGAGCGCCG 1320
Db 421 AspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSerHisLysGlyArg 440
QY 1321 CCGGCAACTTCTTGCAAGACCGCAGCGAGCCCGCCCGCCACCGTGCACCGCCCGCC 1380
Db 441 ProGlyAsnPheLeuGlnSerArgProGlu-----ProThrAlaPro 454
QY 1381 CCGCGCGAGAGCTTCGCTTCGAGAGAGACCAACCCCGCCCGCCCAAGAGAGAGAGAG 1440
Db 455 ProAlaGluSerPheArgPheGluGluThrThrProAlaProLysGlnGluProLysAsp 474
QY 1441 CGGAGACCCCTACCGCGAGCCCTGCAACCGCCCTGCGCAGCGCTGTTCGGCAGCGGCCCTG 1500
Db 475 -----ArgGluProLeuThrSerLeuLysSerLeuPheGlySerAspProLeu 490

QY 1501 AGCCAG 1506
Db 491 SerGln 492

RESULT 11

US-10-339-217-110
; Sequence 110, Application US/10339217
; Publication No. US20030198648A1
; GENERAL INFORMATION:
; APPLICANT: Buechter, Douglas
; APPLICANT: Hou, Xiaohong
; APPLICANT: Marlor, Christopher W.
; APPLICANT: Rice, William G.
; APPLICANT: Yang, Wengang
; TITLE OF INVENTION: Methods for identifying compounds which inhibit binding of
; FILE REFERENCE: 111021.143 (ACH-US1)
; CURRENT APPLICATION NUMBER: US/10/339,217
; PRIOR FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: US 60/347,369
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 110
; LENGTH: 492
; TYPE: PRT
; ORGANISM: HIV-1
US-10-339-217-110

Alignment Scores:
Pred. No.: 6.68e-117 Length: 492
Score: 2430.00 Matches: 458
Percent Similarity: 94.82% Conservative: 18
Best Local Similarity: 91.24% Mismatches: 16
Query Match: 85.56% Indels: 10
DB: 14 Gaps: 2

US-09-475-704A-4 (1-1509) x US-10-339-217-110 (1-492)

QY 1 ATGGGGCCCGCGCCAGCATCTCGCGCGCGAGAAAGTGGACAAGTGGAGAGAATCCGC 60
 Db 1 MetGlyAlaArgAlaSerIleLeuArgGlyGlyLeuAspThrTrpGluIleArg 20

QY 61 CTGCGCCCGCGCGCGAGAAAGCACTAATGCTGAAGCACTGTGTGGCCAGCCGCGAG 120
 Db 21 LeuArgProGlyGlyIleAspCysTyrMetLeuIleValTrpAlaSerArgGlu 40

QY 121 CTGAGGCGCTTCGCGCTGAACCCCGCGCTGTGAGACCGCGAGCGCTGCAAGCAGATC 180
 Db 41 LeuGluArgPheSerLeuAsnProGlyLeuLeuGluThrSerGluGlyCysIleVal 60

QY 181 ATGAAGCAGCTGACGCGCGCGCTGTGAGACCGGACCGAGAGAGCTGCGACCTGTACAAC 240
 Db 61 MetIleGlnLeuGlnProAlaLeuGlnThrGlyThrGluGluLeuIleValAsn 80

QY 241 ACCGTGCGCCACCTGTACTGCGTGACGCGCGCATCGAGTCCGCGACCAAGAGAGGCC 300
 Db 81 ThrValAlaThrLeuPheCysValHisGluIleValAlaArgAspThrIleValAla 100

QY 301 CTGACAAGATCGAGAGAGAGAGACAAGTCCAGCAGAGAAGCCAGAGCGCAAGAGAG 360
 Db 101 LeuAspIleIleGluGluGlnIleAsnIleValSerGlnIleValThrGlnIleValAla 120

QY 361 GCGGAGCGCAAGGTGAGCGCAAGTACCCCATCGTGAGAACTGAGAGCGCGCAGATGGTG 420
 Db 121 AlaAspGlyThrValSerGlnAsnTyrProIleValGlnAsnLeuGlnIleValMetVal 140

QY 421 CACCAAGCCATCAGCGCGCGCACCGCTGAAGCGCTGGTGAAGTGAATGAGAGAGAGGCC 480
 Db 141 HisGlnAlaIleSerProArgThrLeuAsnAlaTrpValIleValIleGluGluValAla 160

QY 481 TTCAGCGCGCGAGGTGATCCCATGTTCACCGCGCTGAGCGAGCGCGCACCGCGCGAGAG 540
 Db 161 PheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGlnAsp 180

QY 541 CTGAACACGATGTTGAACACCGTGGCGCGCACCGAGCGCGCATGAGATGCTGAAGAGAC 600
 Db 181 LeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuIleValAsp 200

QY 601 ACCATCAACGAGAGCGCGCGAGTGGGACCGCGCTGACCGCGCGTGAAGCGCGCGCGCTG 660
 Db 201 ThrIleAsnGluGlnAlaIleGluTrpAspArgLeuHisProValHisAlaGlyProIle 220

QY 661 GCGCGCGCGCGAGATGCGCGCGCGCGCGCGAGCAGATCGCGCGCGCGCACCGAGACCGCTG 720
 Db 221 AlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerThrLeu 240

QY 721 CAGAGCAGATCGCTGATGACCAAGACCGCGCGTGGCGCGCGTGGCGAGCATCTACAAG 780
 Db 241 GlnGluGlnIleAlaTrpMetThrAsnAsnProProValProValGlyAspIleTyrIle 260

QY 781 CGGTGATCATCTCGGCGCTGAACAAGATCGTGGATGTACAGCGCGCTGAGCATCTG 840
 Db 261 ArgTrpIleIleLeuGlyLeuAsnIleValAlaArgMetTyrSerProValSerIleLeu 280

QY 841 GACATCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCCGCGACTACGTGACCGCGCTTCTCAAGACC 900
 Db 281 AspIleArgGlnGlyProIleValProPheArgAspTyrValAlaAspArgPhePheIleThr 300

QY 901 CTGCGCGCGCGAGCAGCGCGCGCGAGAGTGAAGACTGATGACCGAGACCGCTGCTGCTG 960
 Db 301 LeuArgAlaGluGlnAlaThrGlnAspValIleAsnTrpMetThrAspThrLeuLeuVal 320

QY 961 CAGAACGCCCAACCGCGACTGCAAGACCATCTCGCGCTCTCGCGCGCGCGCGCGCGCGCTG 1020
 Db 321 GlnAsnAlaAsnProAspCysIleThrIleLeuArgAlaLeuGlyProGlyAlaThrLeu 340

QY 1021 GAGGAGATGATGACCGCGCTGCGAGGCGGTGGCGCGCGCGCGCGCGCAAGAGCGCGCGTGTG 1080

Db 341 GluGluMetMetThrAlaCysGlnGlyValGlyGlyProSerHisIleValAlaArgValLeu 360

QY 1081 GCCGAGGCGATGAGCCAGCGCGCAACAGCGTGACATCATGATGACAGAAGCACTTCAAG 1140
 Db 361 AlaGluAlaMetSerGlnThrAsnAsnAlaAsnIleMetMetGlnArgSerAsnPheIle 380

QY 1141 GCGCGCGCGCGCGCAAGTCAAGTCTTCACTGCGCGCAAGAGAGGCGCACATCGCGCAAGAAC 1200
 Db 381 GlyProArgArgIleIleIleCysPheAsnIleCysGlyIleGluGlnIleHisLeuAlaArgAsn 400

QY 1201 TGCGCGCGCGCGCGCGCAAGAGGCTGCTGGAAGTGGCGCAAGAGAGGCGCACCATGAAG 1260
 Db 401 CysArgAlaProArgIleValIleCysIleCysTyrIleCysGlyIleGluGlnIleMetIle 420

QY 1261 GACTGCACCGAGCGCGCGCGCAACTTCTCGGCGCAAGATCTGGCGCGCACAGCGCGCGCG 1320
 Db 421 AspCysThrGluArgGlnAlaAsnPheLeuGlyIleTrpProSerHisIleValArg 440

QY 1321 CCGGCGCACTTCTGCGAAGACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
 Db 441 ProGlyAsnPheLeuGlnAsnArgProGlu-----ProThrAlaPro 454

QY 1381 CCGGCGCGAGAGCTTCCGCTTGAAGAGAGACCAACCGCGCGCGCGCGCGCGCGCGCGCG 1440
 Db 455 ProAlaGluSerPheArgPheGluGluThrThrProAlaProIleProIleProArgGlu 474

QY 1441 CCGGAGCGCTTACCGCGAGCGCGCGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCTG 1500
 Db 475 -----ArgGluProLeuThrSerLeuIleValSerLeuPheGlySerAspProLeu 490

QY 1501 AGCCAG 1506
 Db 491 SerGln 492

RESULT 12
 US-10-339-217-140
 ; Sequence 140, Application US/10339217
 ; Publication No. US20030198648A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Buechter, Douglas
 ; APPLICANT: Hou, Xiaohong
 ; APPLICANT: Marlor, Christopher W.
 ; APPLICANT: Rice, William G.
 ; APPLICANT: Yang, Wengang
 ; TITLE OF INVENTION: Methods for Identifying Compounds Which Inhibit Binding of
 ; TITLE OF INVENTION: Nucleocapsid 7 Protein to HIV-1 RNA
 ; FILE REFERENCE: 111021.143(ACH-US1)
 ; CURRENT APPLICATION NUMBER: US/10/339,217
 ; PRIOR FILING DATE: 2003-01-09
 ; PRIOR APPLICATION NUMBER: US 60/347,369
 ; NUMBER OF SEQ ID NOS: 144
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 140
 ; LENGTH: 491
 ; TYPE: PRT
 ; ORGANISM: HIV-1
 US-10-339-217-140

Alignment Scores:
 Pred. No.: 1,83e-116 Length: 491
 Score: 2421.50 Matches: 461
 Percent Similarity: 94.82% Conservative: 15
 Best Local Similarity: 91.83% Mismatches: 15
 Query Match: 85.26% Indels: 11
 DB: 14 Gaps: 3

US-09-475-704A-4 (1-1509) x US-10-339-217-140 (1-491)

QY 1 ATGGGGCCCGCGCCAGCATCTCGCGCGCGAGAAAGTGGACAAGTGGAGAGAATCCGC 60
 Db 1 MetGlyAlaArgAlaSerIleLeuArgGlyGlyLeuAspIleTrpGluIleArg 20

```

QY 61 CTGCCCCCGGCGCAAGACCTACATGCTGAAGCACTGTGTGGGCCAGCCCGGAG 120
Db 21 LeuArgProGlyGlybShiSTyMetLeuYshisLeuValTPrAlaSerArgGlu 40
QY 121 CTGAGAGGCTTCGCCCTGAACCCCGGCTGTGAGACCGCCGAGGCTGCAAGCAGATC 180
Db 41 LeuGluArgPheAlaValAsnProGlyLeuLeuGluThrAlaGluGlyCysbSglnile 60
QY 181 ATGAAGCAGCTGCAGCCCGCTGCAGACCGGCAGAGAGCTGGCAGCCTGTACAAC 240
Db 61 ILeysGlnLeuGlnProAlaLeuGlnThrGlyThrGluGluLeuArgSerLeuYrAsn 80
QY 241 ACCGTGGCCACCTGTACTGCGTGCACGCCGGCATCGAGTCCGCGACCAAGAGGCC 300
Db 81 ThrValAlaThrLeuYrCysValHisAlaGluIleGluValArgAspThrlySglnAla 100
QY 301 CTGACAAGATCGAGAGAGAGAGACAAGTCCAGCAGAAAGACCGAGCCCAAGAG 360
Db 101 LeuAspArgIleGluGlnGluGlnAsnLysSerGlnLysThrGlnGlnAlaAsnGlu 120
QY 361 GCCGACGGCAAGGTGAGCCAGAACTACCCCATCTGTGAGAACTTGACAGGCGCAGATG 420
Db 121 AlaAspGlyLysValSerGlnAsnYrProIleValGlnAsnLeuGlnGlyGlnMetVal 140
QY 421 CACCAAGCCATCAGCCCCCGCACCTCTGAACGCTGGGTGAAGTGAATCGAGAGAGGCC 480
Db 141 HisGlnAlaIleSerProArgThrLeuAsnAlaTPrValLysValIleGluGluYsAla 160
QY 481 TTCAGCCCCGAGGTGATCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCGCAGAC 540
Db 161 PheSerProGluValIleProMetPheThrAlaLeuSerGluGlyAlaThrProGlnAsp 180
QY 541 CTGAACAAGATGTTGAACACCGTGGCGCGCCACAGAGCCCGCATGAGATGCTGAAGAC 600
Db 181 LeuAsnThrMetLeuAsnThrValGlyGlnHisGlnAlaAlaMetGlnMetLeuYsAsp 200
QY 601 ACCATCAACGAGAGAGCGCGCAGTGGAGCCGCTGTCAACCCCGTGACAGCGCGCCGCTG 660
Db 201 ThrIleAsnGluGlnAlaAlaGluTPrAspArgLeuHisProValHisAlaGlyProIle 220
QY 661 GCCCGCGCGCAGATGCGCGCAGACCCCGCGGCGAGCGACATCGCGCGCCACAGACCCCTG 720
Db 221 AlaProGlyGlnMetCArgGluProArgGlySerAspIleAlaGlyThrThrSerSerLeu 240
QY 721 CAGAGCAGATCGCCTGGATGACCAACACCCCGCTGCCGCTGGCGGCGACATCTTACAAG 780
Db 241 GlnGluGlnIleAlaTPrMetThrGlyAsnProProValProValGlyAspIleYrlys 260
QY 781 CGGTGATCATCCTGGGCTTGAACAAGATCGTGGGATGTACAGCCCGGTGAGCATCTCG 840
Db 261 ArgTPrIleIleLeuGlyLeuAsnLysIleValArgMetYrSerProValSerIleLeu 280
QY 841 GACATCCGCGCAGCGCCCAAGAGACCCCTTCGCGGACTACGTGACCGCTTCTTCAAGAC 900
Db 281 AspIleLysGlnGlyProLysGluProPheArgAspTyrValAspArgPhePheLysThr 300
QY 901 CTGGCGCGCGAGCAGGCCACCCAGACGTGAAGACTGATGACCGAGACCCCTGCTGTG 960
Db 301 LeuArgAlaGluGlnAlaThrGlnAspValLysAsnTPrMetThrAspThrLeuLeuVal 320
QY 961 CAGAACGCCAACCCCGACTGCAAGACCATCTCGCGGCTCTCGGCGCCCGGCGCACCTG 1020
Db 321 GlnAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeuGlyProGlyAlaSerLeu 340
QY 1021 GAGGAGATGATGACCGCTGCCAGGCGTGGGCGCGCCCGGCGCAAGGCCCGCGTCTG 1080
Db 341 GluGlnMetMetThrAlaCysGlnGlyValGlyLysProSerHisLysAlaArgValLeu 360
QY 1081 GCCGAGCGCATGAGCCAGGCCAACAGCGTGAACATCATGATGACAGAAAGCAACTTCAAG 1140
Db 361 AlaGlnAlaMetSerGlnThrAsnSer--ThrIleLeuMetGlnArgSerAsnPhelys 379
QY 1141 GGCCCCCGCGCAAGCTCAAGTGTCTTCACTGCGGCAAGAGGGCCACATCGCCAAAGAAC 1200

```

```

Db 380 GlyProLysArgIleValLysCysPheAsnYsGlyLysGlnGlyHisIleAlaLysAsn 399
QY 1201 TGC CGCGCGCGCGCGCAAGAGGCTGCTGGAAGTCCGCGGCAAGAGAGGCCACAGATGAAG 1260
Db 400 CysArgAlaProArgLysLysGlyCysTPrLysCysGlyLysGlnGlyHisGlnMetLys 419
QY 1261 GACTGCACCGAGCGCCAGGCCCACTTCTGGGCAAGATCTGGGCCCAAGCCAAAGGGCCG 1320
Db 420 AspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTPrProSerHisLysGlyArg 439
QY 1321 CCGGCGCACTTCTTGCAAGAACCGCGCAGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Db 440 ProGlyAsnPheLeuGlnSerArgProGlu-----ProThrAlaPro 453
QY 1381 CCGCGCGAGAGCTTCGCTTCGAGAGAGACCAACCCCGCGCGCAAGAGAGCCCAAGAG 1440
Db 454 ProAlaGlnSerPheArgPheGlnGluThrThrProAlaLeuGlnGlnGlyProLysAsp 473
QY 1441 CGCGAGCCCTACCGCGAGCGCGCTGACCGCGCTGCGCAGCGCTGTTCGGCAGCGCGCTG 1500
Db 474 -----ArgGluProLeuThrSerLeuArgSerLeuPheGlySerAspProLeu 489
QY 1501 AGCCAG 1506
Db 490 SerGln 491

RESULT 13
US-10-339-217-143
; Sequence 143, Application US/10339217
; Publication No. US20030198648A1
; GENERAL INFORMATION:
; APPLICANT: Buechter, Douglas
; APPLICANT: Hou, Xiaohong
; APPLICANT: Marlor, Christopher W.
; APPLICANT: Rice, William G.
; APPLICANT: Yang, Wengang
; TITLE OF INVENTION: Methods for Identifying Compounds Which Inhibit Binding of
; TITLE OF INVENTION: Nucleocapsid 7 Protein to HIV-1 RNA
; FILE REFERENCE: 111021.143(ACH-US1)
; CURRENT APPLICATION NUMBER: US/10/339,217
; CURRENT FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: US 60/347,369
; PRIOR FILING DATE: 2002-01-11
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 143
; LENGTH: 491
; TYPE: PRT
; ORGANISM: HIV-1
US-10-339-217-143

Alignment Scores:
Pred. No.: 4.7e-116 Length: 491
Score: 2413.50 Matches: 456
Percent Similarity: 95.02% Conservative: 21
Best Local Similarity: 90.84% Mismatches: 14
Query Match: 84.98% Indels: 11
DB: 14 Gaps: 3

US-09-475-704A-4 (1-1509) x US-10-339-217-143 (1-491)
QY 1 ATGGGCGCGCGCGCGCAGCATCTTCGCGCGGAGAGAGCTGGAACAAGTGGAGAAATCCGC 60
Db 1 MetGlyAlaArgAlaSerIleLeuArgGlyGlyLysLeuAspLysTPrGluLysIleArg 20
QY 61 CTGGCGCGCGCGGCAAGAGCACTACATCTGAAGCAGCCTGTGTGGGCGCAGCGCGAG 120
Db 21 LeuArgProGlyGlyLysLysHisTyrMetIleLysHisLeuValTPrAlaSerArgGlu 40
QY 121 CTGAGAGGCTTCGCCCTGAACCCCGGCTGTGAGACCGCCGAGGCTGCAAGCAGATC 180
Db 41 LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrAlaGluGlyCysbSglnile 60

```



```

OY 181 ATGAGCAGCTGCAGCCCGCCCTGAGACCGGACAGAGAGCTGGCAGCCTGTACAAC 240
Db 61 ILeysGlnLeuGlnProAlaLeuGlnThrGlnGluLeuArgSerLeuHisAsn 80
OY 241 ACCGTGGCCACCCTGTACTGCGTGCACGCGGCATGAGGTCCGCGACCAAGAGGCC 300
Db 81 ThrValAlaThrLeuTyrCysValHisAlaGlyIleGlnIleArgAspThrLysGluAla 100
OY 301 CTGGAACAAGATCGAGAGAGAGACAGAACAAGTCCCGACAGAGAGCCAGAGAGAG 360
Db 101 LeuAspLysIleGlnGlnGlnGlnGlnLysSerGlnGlnLysThrGlnGlnAlaLysGlu 120
OY 361 GCCGAGCGCAAGGTGAGCGCAGAACTAACCTATCTGTGAGAACTGTGAGGGCCAGATG 420
Db 121 AlaAspGlyLysValSerGlnAsnTyrProIleValGlnAsnLeuGlnGlyGlnMetVal 140
OY 421 CACCAAGGCATCAGCCCGCCGACACCTGTAAAGCCTGGGTGAAGTGAATCGAGAGAGGCC 480
Db 141 HisGlnAlaLeuSerProArgThrLeuAsnAlaTyrValLysValIleGlnGluLysAla 160
OY 481 TTCAAGCCCGAGGTGATCCCATATGTTCAACGCGCTGAGCGAGGGCGCCACCCCGAGAC 540
Db 161 PheSerProGlnIleIleProMetPheThrAlaLeuSerGlnGlyAlaThrProGlnAsp 180
OY 541 CTGAACAAGATGTTGAACAACCGTGGCGCGCCACAGAGCCGCGCATGATGCTGAAGAC 600
Db 181 LeuAsnThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLysAsp 200
OY 601 ACCATCAACGAGAGAGCGCGCGAGTGGGAGCGGCTGTCAACCCCGTGCAGCGCGCCGCTG 660
Db 201 ThrIleAsnGlnGluAlaAlaGlnTyrAspArgLysHisProAlaGlnAlaGlyProIle 220
OY 661 GCCCGCGCGCAGATGCGCGCGACCCCGCGCGAGCGAGCATGCGCGCGCCACAGACCCCTG 720
Db 221 AlaProGlyGlnMetArgGluProArgGlySerAspIleAlaGlyThrThrSerThrLeu 240
OY 721 CAGGAGCAGATGCGCTGATGACCAACACCCCGCTGCGCGCGAGCATTTACAAG 780
Db 241 GlnGlnGlnIleAlaTyrMetThrGlyAsnProProValProValGlyGluIleTyrLys 260
OY 781 CGGTGGATCATCTTGAGGCTGAAACAAGATCGTGGATGTACAGCCCGGTGAGCATCTTG 840
Db 261 ArgTyrPheIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProValSerIleLeu 280
OY 841 GACATCCGCGCAGGCGCGCCCAAGAGAGCGCTTCCGCGACTACGTGACCGCTTTTCAAGAC 900
Db 281 AspIleLysGlnGlyProLysGluProPheArgAspTyrValAspArgPhePheLysThr 300
OY 901 CTGCGCGCGCAGAGAGCGCACCCAGAGAGCTGAAGAAGTGAAGACCGAGACCCCTGCTG 960
Db 301 LeuArgAlaGlnGlnAlaThrGlnAspValLysAsnTyrMetThrAspThrLeuLeuVal 320
OY 961 CAGAACGCCAACCCCGACTGCAGAACCATCTGCGCGCTCTCGCGCGCGCGACCCCTG 1020
Db 321 GlnAsnAlaAsnProAspCysLysThrIleLeuArgAlaLeuGlyProGlyAlaSerLeu 340
OY 1021 GAGGAGATGATGACCGCTGCGCAGGCGTGGCGCGCCCGGACAGAGCCCGCGTGTG 1080
Db 341 GlnGlnMetMetThrAlaCysGlnGlyValGlyLysProSerHisLysAlaArgValLeu 360
OY 1081 GCCGAGCGCATGAGCGCAGGCCAACAGCGTGAACATCATGATGCAGAGAAGCAACTTCAAG 1140
Db 361 AlaGlnAlaMetSerGlnThrAsnAsn--SerIleLeuMetGlnArgSerAsnPheLys 379
OY 1141 GGCCCCCGCGCAAGCTCAAGTGTCTCACTGCGGCAAGAGAGGCCACATCGCCAAGAAC 1200
Db 380 GlyPheLysArgThrValLysCysPheAsnCysGlyLysGlnGlyHisIleAlaArgAsn 399
OY 1201 TGCGCGCGCGCGCGCAAGAAGGCGTCTGGAAGTGCAGGCAAGAGGCCACCAAGTGAAG 1260
Db 400 CysArgAlaProArgLysLysGlyCysTyrLysCysGlyLysGlnGlyHisGlnMetLys 419

```

```

OY 1261 GACTGCACCGAGCGCCAGGCCCACTTCCTGGGCAAGATCTTGCCCCAGCCACAGAGGCCGC 1320
Db 420 AspCysThrGlnArgGlnAlaAsnPheLeuGlyLysIleTyrProSerHisLysGlyArg 439
OY 1321 CCGCGCAACTTCCTTGCAAGAACCGCAGCGAGCGCGCGCGCCCGCCACCGTGCACCGCCCC 1380
Db 440 ProGlyAsnPheLeuGlnAlaAsnArgProGlu-----ProThrAlaPro 453
OY 1381 CCGCGCGAGAGCTTCGCTTGCAGAGAGACCAACCCCGCGCGCGCGCGCGCGCGCGCGCG 1440
Db 454 ProAlaGlnSerPheArgPheGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 473
OY 1441 CCGGAGCCCTTACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500
Db 474 -----ArgGluProLeuThrSerLeuLysSerLeuPheGlySerAspProLeu 489
OY 1501 AGCCAG 1506
Db 490 SerGln 491

```

RESULT 14

```

US-10-339-217-105
; Sequence 105, Application US/10339217
; Publication No. US20030198648A1
; GENERAL INFORMATION:
; APPLICANT: Buechter, Douglas
; APPLICANT: Hou, Xiaohong
; APPLICANT: Marlor, Christopher W.
; APPLICANT: Rice, William G.
; APPLICANT: Yang, Wengang
; TITLE OF INVENTION: Methods for Identifying Compounds Which Inhibit Binding of
; FILE REFERENCE: 111021.143(ACH-US1)
; CURRENT APPLICATION NUMBER: US/10/339,217
; PRIOR FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: US 60/347,369
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 105
; LENGTH: 508
; TYPE: PRT
; ORGANISM: HIV-1
US-10-339-217-105

```

```

Alignment Scores:
Pred. No.: 1,08e-114 Length: 508
Score: 2387.00 Matches: 454
Percent Similarity: 92.19% Conservative: 18
Best Local Similarity: 88.67% Mismatches: 26
Query Match: 84.05% Indels: 14
DB: 14 Gaps: 3

```

US-09-475-704A-4 (1-1509) x US-10-339-217-105 (1-508)

```

OY 1 ATGGGGCGCGCGCCAGCATCTGCGCGGAGAGAGTGGACAAGTGGAGAAATCCGC 60
Db 1 MetGlyAlaArgAlaSerIleLeuArgGlyThrLysLeuAspAlaTyrGlnLysIleArg 20
OY 61 CTGCGCGCGCGCGCGCAAGAGCACTACATGCTGAAGCACTGTGTGTGGCGCGCGAG 120
Db 21 LeuArgProGlyGlyLysLysHisTyrMetLeuLysHisIleValTyrAlaSerArgGlu 40
OY 121 CTGAGAGGCTTCGCGCTGAACCCCGCGCTGTGAGAGCCGCGAGAGGCTGCAAGCATC 180
Db 41 LeuGlnArgPheAlaLeuAsnProGlyLeuLeuGlnThrSerGlnGlyCysLysGlnIle 60
OY 181 ATGAAGCAGCTGCAGCGCGCGCTGCAGACCGGCAACGAGAGCTGCGCAGCTGTACAAC 240
Db 61 MetLysGlnLeuHisProAlaLeuGlnThrGlyThrGlnGlnLeuLysSerLeuTyrAsn 80
OY 241 ACCGTGGCCACCTGTACTGTGCTGCACCGCGGATCGAGGTCCGCGACACCAAGAGGCC 300

```

```

Db      81 ThrValAlaThrLeuTyrCysValHisGluAsnIleIysValArgAspThrLysGluAla 100
QY      301 CTGACACAGATCGAGAGAGACAGAAC-----AAGTCCACAGAGAACCCAGCAGGCC 354
Db      101 LeuAspLysIleGluGluGluGlnAsnLysIleLysSerGlnGlnLysThrGlnGlnAla 120
QY      355 AAGAGCGCCGACGGCAGAGTGAGCCAGAACTACCCCATGTGCAGAACCTGCAGGGGCCAG 414
Db      121 LysAlaAlaAspGluLysValSerGlnAsnTyrProIleValGlnAsnLeuGlnGln 140
QY      415 ATGTGACACAGGCCATCAGCCCCCGCACCTTGAAACGCTGGGTGAAGTGATCGAGAG 474
Db      141 MetValHisGlnAsnLeuSerProArgThrLeuAsnAlaTrpValLysValIleGluGlu 160
QY      475 AAGGCTTTCAGCCCCGAGGTGATCCCCCATGTTCACCGCCTGAGCGAGGGCCACCCCC 534
Db      161 LysAlaPheSerProGluValIleProMetPheThrAlaLeuSerGlnGlyAlaThrPro 180
QY      535 CAGGACCTGAACACGATGTTGAACACCGTGGGGCGGCACAGCGCCGACATGCAGATGCTG 594
Db      181 GlnAspLeuSerThrMetLeuAsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeu 200
QY      595 AAGGACACCATCAACGAGAGAGCGCCGAGTGGAGCCGCTGCACCCCGTGCAGAGCCGCG 654
Db      201 LysAspThrIleAsnGluGlnAlaAlaGluTrpAspArgLeuHisProValHisAlaGly 220
QY      655 CCGGTGGCCCCCGGCAGATGCGGACCCCGCGCAGCAGACATCGCCGCGCCACACAGC 714
Db      221 PrometAlaProGlyGlnLeuArgGluProArgGlySerAspIleAlaGlyThrThrSer 240
QY      715 ACCCTGCAGAGACAGATCGCCTGTGATGACAGCAACCCCGCTGCAGCGTGGCGACATC 774
Db      241 ThrLeuArgGluGlnIleAlaThrMetThrSerAsnProProIleProValGlyAspIle 260
QY      775 TACAAGCGGTGATCATCTGTGGCCTGAACAAGATCGTGGATGTACAGCCCGTGAGC 834
Db      261 TyrLysArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProValSer 280
QY      835 ATCTTGACATCCGCGCAGAGCGCCCAAGAGCCCTTCCGCACTACGTGAGCCGTTTTC 894
Db      281 IleLeuAspIleArgGlnGlyProLysGluProPheArgAspTyrValAspArgPhe 300
QY      895 AAGACCTGCGCGCGCAGAGCGCCACCGACGAGAGTGAAGAACTGATGACCGAGACCTG 954
Db      301 LysAlaLeuArgAlaGluGlnAlaThrGlnAspValLysAsnTrpMetThrGluThrLeu 320
QY      955 CTGCTGAGAACGCGCAACCCCGACTGCAAGACCATCTCGCGCTCGGCCCGCGGCC 1014
Db      321 LeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyIleGlyAla 340
QY      1015 ACCCTGAGAGATGATGACCGCCTGCAGAGGGCGTGGGGGCCCGCCACAGGCCCGC 1074
Db      341 ThrLeuGluGluMetMetThrAlaCysGlnGlyValGlyGlyProSerHisLysAlaArg 360
QY      1075 GTGCTGGCCGAGCGCAGTGAAGCCAGGCCAACAGCGTGAACATCATGTATGACAGAGCAAC 1134
Db      361 ValLeuAlaGluAlaMetSerGlnAlaAsnAsnThrAsnIleMetMetGlnArgSerAsn 380
QY      1135 TTCAAGGGCCCCCGCGCAACGTCAAGTCTTCACTGCGGCAAGAGGGCCACATCGCC 1194
Db      381 PheLysSerSerLysArgIleValLysCysSerAsnLysGlyLysGluGlyHisIleAla 400
QY      1195 AAGAACTGCGCGCGCCCCCGCAAGAGAGGCTGCTGAAGTGGCGCAAGAGGGCCACAG 1254
Db      401 ArgAsnCysArgAlaProArgLysLysGlyCysTrpLysCysGlyLysGluGlyHisGln 420
QY      1255 ATGAAGACTGCACCGCAGCGCCACCACTTCTGGGCAAGATCTGGCCCAAGCCACAG 1314
Db      421 MetLysAspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSerHisLys 440
QY      1315 GGCCGCGCGCACTTCTCTGCAAGACCGCAGCGCCCGCCCGCCACCGTG----- 1368
Db      441 GlyArgProGlyAsnPheLeuGlnAsnArgProGluProThrAlaProAlaGluSer 460

```

```

QY      1369 -----CCACCGCCCCCGCGCAGAGCTTCGCTTCGAGAGACC 1410
Db      461 PheArgAsnArgProGluProThrAlaProProAlaGluSerPheArgPheGluGluThr 480
QY      1411 ACCCGCGCCCCCAAGAGAGAGCCCAAGACCGCGACCCCTACCGCGAGCCCTGACCGCC 1470
Db      481 ThrProThrProLysGlnGluProLysAsp-----ArgAspProLeuThrSer 496
QY      1471 CTGCGAGCCTGTTCGCGACGCGGCCCTGAGCCAG 1506
Db      497 LeuLysSerLeuPheGlySerAspProSerSerGln 508

RESULT 15
US-10-339-217-104
; Sequence 104, Application US/10339217
; Publication No. US20030198648A1
; GENERAL INFORMATION:
; APPLICANT: Buechter, Douglas
; APPLICANT: Hou, Xiaohong
; APPLICANT: Marlor, Christopher W.
; APPLICANT: Rice, William G.
; APPLICANT: Yang, Wengang
; TITLE OF INVENTION: Methods for Identifying Compounds which Inhibit Binding of
; TITLE OF INVENTION: Nucleocapsid 7 Protein to HIV-1 RNA
; FILE REFERENCE: 111021.143(ACH-US1)
; CURRENT APPLICATION NUMBER: US/10/339,217
; PRIOR FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: US 60/347,369
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 104
; LENGTH: 491
; TYPE: PRT
; ORGANISM: HIV-1
US-10-339-217-104

Alignment Scores:
Pred. No.: 1.29e-114 Length: 491
Score: 2385.50 Matches: 453
Percent Similarity: 93.82% Conservative: 18
Best Local Similarity: 90.24% Mismatches: 20
Query Match: 84.00% Indels: 11
DB: 14 Gaps: 3

US-09-475-704A-4 (1-1509) x US-10-339-217-104 (1-491)
QY      1 ATGGGCGCGCGCGCGCAGCATCTCGGCGCGCGAGAGCTGGAACAGTGGAGAGATCCGC 60
Db      1 MetGlyAlaArgAlaSerIleLeuArgGlyGluLysLeuAspLysTrpGluLysIleArg 20
QY      61 CTGCGCGCGCGCGCGCAAGACACTACATGCTGAAGCACTGTGTGTGGCCAGCGCGGAG 120
Db      21 LeuArgProGlyGlyArgLysHisTyrMetLeuLysHisLeuValTrpAlaSerArgGlu 40
QY      121 CTGAGGGCTTGGCCTTGAACCCCGCCTGTGTGAGACCGCCGAGGGCTGCAAGCAGATC 180
Db      41 LeuGluArgPheAlaLeuAsnProGlyLeuLeuGluThrSerGlnGlyCysLysGlnIle 60
QY      181 ATGAAGCAGCTGCAGCGCCCGCCTGCAAGACCGGCAACGAGAGCTGCGCAGCCTGTACAAC 240
Db      61 IleLysGlnLeuHisProAlaLeuLysThrGlyThrGluGluLeuArgSerLeuTyrAsn 80
QY      241 ACCGTGGCCACCTGTACTGTGCGTGCACGCGCGGATCGAGGTCCGCGACACCAAGAGGCC 300
Db      81 ThrValAlaThrLeuTyrCysValHisGluAsnIleGluValArgAspThrLysGluAla 100
QY      301 CTGACCAAGATCGAGAGAGAGAGCAACAAGTCCCAAGCAAGAACCCAGAGGCCAAGAG 360
Db      101 LeuAspLysIleGluGluGluGlnAsnLysSerGlnGlnLysThrGlnGlnAlaLysAla 120
QY      361 GCCGACGGCAAGGTGAGCCAGAACTACCCCATGTGCAAGAACTGCAAGAGGGCCAGATGTG 420

```

```
Db      121  ALaaspGluGlyValSerGlnAsnTyrProIleValGlnAsnLeuGlnGlyGlnMetVal 140
QY      421  CACCAAGCCATCAGCCCCCGCACCCCTGAACGCCCTGGGTGAAGTGAATCGAGAGAGAGCC 480
Db      141  HIsGlnAlaIleSerProArgThrLeuAsnAlaTrpValIleGluGluIysAla 160
QY      481  TTCAGCCCCGAGGTGATCCCATGTTCAACGCCCTGAGCGAGGCGCCACCCCGCAGAC 540
Db      161  PheSerProGluValIlePrometPheThrAlaLeuSerGluGlyAlaThrProGlnAsp 180
QY      541  CTGAACAAGATGTTGAACAACCGTGGCGGCCACCAAGCCCATGAGATGCTGAAGAC 600
Db      181  LeuAsnThrMetLeuAsnThrValGlyGlnHisGlnAlaAlaMetGlnMetLeuIysAsp 200
QY      601  ACCATCAACGAGAGGCGCGCGAGTGGGACCGCCTGCAACCCCGTGACGCGCGCCGCTG 660
Db      201  ThrIleAsnGluGluAlaAlaGluTrpAspArgLeuHisProValHisAlaGlyProAla 220
QY      661  GCCCGCGCCGAGATGCGCGCACCCCGCGGCGACGACATCGCCGCGGCCACCAACCTTG 720
Db      221  AlaProGlyGlnMetCArgGluProArgGlySerAspIleAlaGlyThrThrSerThrLeu 240
QY      721  CAGAGCAGATCGCCTGGATGACCAACACCCCGCTGCGCGCGGCGACATCTACAAG 780
Db      241  GlnGluGlnIleAlaTrpMetThrGlyAsnProProValProValGlyAspIleTyrLys 260
QY      781  CGGTGATCATCCTGGGCTTGAACAAGATCGTGCGGATGTACAGCCCGTGAGCATCCTG 840
Db      261  ArgTrpIleIleLeuGlyLeuAsnLysIleValArgMetTyrSerProValSerIleLeu 280
QY      841  GACATCCGCGCAGGCGCCCAAGAGAGCCCTTCGCGACTACGTGACCGCTTCTCAAGACC 900
Db      281  AspIleLysGlnGlyProLysGluProPheArgAspTyrValAspArgPhePheLysVal 300
QY      901  CTGCGCGCCGAGCAGGCCACCCAGAGACGTGAAGAAGTGAATGATGACCGAGACCTGTG 960
Db      301  LeuArgAlaGluGlnAlaThrGlnAspValLysAsnTrpMetThrAspThrLeuLeuIle 320
QY      961  CAGAACGCCAACCCGCACTGCAAGACCATCTGCGCGCTTCGCGCCGCGGCCACCTTG 1020
Db      321  GlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGlyProAlaAlaSerLeu 340
QY      1021  GAGGAGATGATGACCGGCTGCCAGGGGCTGGCGGCCCGGCCACAGGCCCGCTGCTG 1080
Db      341  GluGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisIlybAlaArgValLeu 360
QY      1081  GCCGAGCGGATGAGCGCGCCCAACAGCGTGAACATCATGATGACAGAAGCACTTCAAG 1140
Db      361  AlaGlnAlaMetSerGlnAlaAsnSer---AsnIleMetMetGlnArgSerAsnPheLys 379
QY      1141  GGGCCCCGGCGCAAGTCAAGTGTCAACTGCGGCAAGAGGGGCCACATCGCCCAAGAAC 1200
Db      380  GlySerLysArgIleValLysCysPheAsnCysGlyLysGluGlyHisIleAlaArgAsn 399
QY      1201  TGCCGGGCCCCCGCAAGAAGGGCTGTGGAAGTGGCGCAAGAGGGCCACCAAGATGAAG 1260
Db      400  CysArgAlaProArgLysLysGlyCysTrpLysCysGlyGlnGlyHisGlnMetLys 419
QY      1261  GACTGCAACCGAGCGCCAGGCCAATCTCTGGGCAAGATCTGGCCCAAGAGGGCGCG 1320
Db      420  AspCysThrGluArgGlnAlaAsnPheLeuGlyLysIleTrpProSerHisLysGlyArg 439
QY      1321  CCGGCAACTTCTCTGAGAACCAGCGAGGAGCCCGCGCCCAACCTGCCCCACCGCCCC 1380
Db      440  ProGlyAsnPheLeuGlnSerArgProGlu-----ProThrAlaPro 453
QY      1381  CCGCGCAGAGAGCTTCCGCTTGAAGAGACCAACCCCGCCCCCAAGCAGAGAGCCCAAG 1440
Db      454  ProAlaGluSerPheArgPheGlnGluThrThrProAlaProLysGlnGluSerLysAsp 473
QY      1441  CGCGAGCCCTACCGCGAGCCCTGACCGCCCTGCGCAGCCTGTTCCGCAAGCGGCCCTG 1500
```

```
Db      474  -----ArgGluProLeuIleSerLeuLysSerLeuPheGlySerAspProSer 489
QY      1501  AGCCAG 1506
Db      490  SerGln 491
```

Search completed: September 1, 2005, 23:51:13
Job time : 408.088 secs